

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 14, 2002, 04:24:33 ; Search time 75.1 seconds  
(without alignments)  
6308.815 Million cell updates/sec

Title: US-09-646-825-1

Perfect score: 2092

Sequence: 1 gaattctctagactccacca.....cttggttaagagctcaagctt 2092

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 351203 seqs, 113238999 residues

Total number of hits satisfying chosen parameters: 702406

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

Issued\_Patents\_NA: \*  
1: /cgn2\_6/ptodata/1/ina/5A\_COMB.seq: \*  
2: /cgn2\_6/ptodata/1/ina/5B\_COMB.seq: \*  
3: /cgn2\_6/ptodata/1/ina/6A\_COMB.seq: \*  
4: /cgn2\_6/ptodata/1/ina/6B\_COMB.seq: \*  
5: /cgn2\_6/ptodata/1/ina/PCTUS\_COMB.seq: \*  
6: /cgn2\_6/ptodata/1/ina/backfiles1.seq: \*

-Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Query	Score	Match	Length	DB	ID	Description
1	147.6	7.1	694	4	US-08-998-416-741		Sequence 741, App
2	90.2	4.3	711	4	US-08-998-416-633		Sequence 633, App
3	48.6	2.3	7218	1	US-08-232-463-14		Sequence 14, App
4	39.2	1.9	7218	1	US-08-232-463-14		Sequence 14, App
5	35.6	1.7	964	1	US-07-798-223A-1		Sequence 1, Appl
6	34	1.6	4724	1	US-08-404-665-3		Sequence 3, Appl
7	34	1.6	4724	1	US-08-404-671-3		Sequence 3, Appl
8	34	1.6	4724	1	US-08-404-781-3		Sequence 3, Appl
9	33.6	1.6	4450	3	US-08-617-860B-2		Sequence 3, Appl
10	33.6	1.6	4706	2	US-08-331-081B-5		Sequence 5, Appl
11	33.6	1.6	4744	2	US-08-331-081B-3		Sequence 3, Appl
12	32.8	1.6	1293	1	US-08-476-008-43		Sequence 43, Appl
13	32.8	1.6	1293	1	US-08-306-063-43		Sequence 43, Appl
14	32.8	1.6	1293	1	US-08-833-485-43		Sequence 43, Appl
15	32.8	1.6	1293	4	US-09-137-440-43		Sequence 43, Appl
16	32	1.5	5191	1	US-08-340-428B-1		Sequence 1, Appl
17	32	1.5	5191	5	PCT-US93-07306-1		Sequence 1, Appl
18	31.8	1.5	1508	3	US-09-041-991A-3		Sequence 3, Appl
19	31.6	1.5	246240	2	US-08-724-394A-20		Sequence 20, Appl
20	31.6	1.5	246240	2	US-08-724-394A-21		Sequence 21, Appl
21	31.6	1.5	246240	2	US-08-724-394A-22		Sequence 22, Appl
22	31	1.5	432	4	US-09-008-697A-19		Sequence 19, Appl
23	30.6	1.5	289	4	US-09-007-005-17		Sequence 17, Appl
24	30.6	1.5	289	4	US-09-244-796-17		Sequence 17, Appl
25	30.6	1.5	3656	1	US-08-232-463-8		Sequence 8, Appl
26	30.6	1.5	9454	1	US-08-232-463-3		Sequence 3, Appl
27	30.6	1.5	9454	1	US-08-232-463-4		Sequence 4, Appl

28 30.2 1.4 573 1 US-08-102-757-20 Sequence 20, Appl  
29 30.2 1.4 1645 5 PCT-US94-01321-9 Sequence 9, Appl  
30 30 1.4 1083 2 US-09-057-762-24 Sequence 24, Appl  
31 30 1.4 1083 3 US-08-326-119A-24 Sequence 24, Appl  
32 30 1.4 2364 2 US-08-838-219B-5 Sequence 5, Appl  
33 30 1.4 2364 3 US-09-233-336A-5 Sequence 5, Appl  
34 30 1.4 2364 3 US-09-233-752A-5 Sequence 5, Appl  
35 30 1.4 2364 4 US-09-402-036-5 Sequence 5, Appl  
36 30 1.4 2375 4 US-08-960-780-3 Sequence 3, Appl  
37 30 1.4 2375 4 US-08-960-780-7 Sequence 7, Appl  
38 30 1.4 2375 4 US-09-073-898-3 Sequence 3, Appl  
39 30 1.4 2375 4 US-09-073-898-7 Sequence 3, Appl  
40 30 1.4 2378 1 US-08-471-033-28 Sequence 28, Appl  
41 30 1.4 2378 2 US-08-471-044-28 Sequence 28, Appl  
42 30 1.4 2378 2 US-08-463-483A-28 Sequence 28, Appl  
43 30 1.4 2378 2 US-08-471-046A-28 Sequence 28, Appl  
44 30 1.4 2378 2 US-08-470-566B-28 Sequence 28, Appl  
45 30 1.4 2378 2 US-08-838-219B-1 Sequence 1, Appl

#### ALIGNMENTS

RESULT 1  
US-08-998-416-741  
; Sequence 741, Application US/08998416  
; Patent No. 6239264  
; GENERAL INFORMATION:  
; APPLICANT: Philippsen, Peter  
; APPLICANT: Pohlmann, Rainer  
; APPLICANT: Steiner, Sabine  
; APPLICANT: Mohr, Christine  
; APPLICANT: Wendland, Jurgen  
; APPLICANT: Knechtel, Philipp  
; APPLICANT: Reibschung, Corinne  
; TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBYA GOSSYPPII  
; NUMBER OF SEQUENCES: 1152  
; CORRESPONDENCE ADDRESS:  
; ADDRESSES: No. 6239264artis Corporation  
; STREET: 3054 Cornwallis Road  
; CITY: Research Triangle Park  
; STATE: No. 6239264th Carolina  
; COUNTRY: USA  
; ZIP: 27709  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/998,416  
; FILING DATE: 24-DEC-1997  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: CH 0016/97  
; FILING DATE: 31-DEC-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Meigs, J. Timothy  
; REGISTRATION NUMBER: 38,241  
; REFERENCE/DOCKET NUMBER: PF/5-30305/A/CGC1976  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 919-541-8587  
; TELEFAX: 919-541-8689  
; INFORMATION FOR SEQ ID NO: 741:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 694 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; ORIGINAL SOURCE:  
; ORGANISM: PAG1482RP











LENGTH: 4744 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-331-081B-3

Query Match 1.6%; Score 33.6; DB 2; Length 4744;  
Best Local Similarity 55.0%; Pred. NO. 2.2;  
Matches 66; Conservative 0; Mismatches 54; Indels 0; Gaps 0;  
QY 364 agtttttccaaacggttgatggcaaatgagacgacctatcactactactatgaggaataa 423  
DB 3448 AGTTATTACTAACCAATCTTGCATAGAGAACACAAAAAACAATGCACCTGTAGAAAA 3389  
QY 424 ctatggatccactgaattgatcgatctcaatgggtgcgcatgggctcgctcttt 483  
DB 3388 TACAGGAGGCACATGAATTGAAGGCAACATGGTTGGCACCTTGTGCATGCTCTCTT 3329

RESULT 12  
US-08-476-008-43  
Sequence 43, Application US/08476008  
Patent No. 5627061  
GENERAL INFORMATION:  
APPLICANT: Barry, Gerard F.  
APPLICANT: Kishore, Ganesh M.  
APPLICANT: Padgett, Stephen R.  
APPLICANT: Stallings, William C.  
TITLE OF INVENTION: Glyphosate Tolerant  
NUMBER OF SEQUENCES: 69  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Dennis R. Hoerner, Jr., Monsanto Co. BB4F  
STREET: 700 Chesterfield Village Parkway  
CITY: St. Louis  
STATE: Missouri  
COUNTRY: USA  
ZIP: 63198  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/476,008  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/306,063  
FILING DATE: 13-SEP-1994  
APPLICATION NUMBER: US 07/749,611  
FILING DATE: 28-AUG-1991  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/576,537  
FILING DATE: 31-AUG-1990  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Hoerner Jr., Dennis R.  
REGISTRATION NUMBER: 30,914  
REFERENCE/DOCKET NUMBER: 38-21(10660)A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (314)537-6099  
TELEFAX: (314)537-6047  
INFORMATION FOR SEQ ID NO: 43:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1293 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
FEATURE:

NAME/KEY: CDS  
LOCATION: 1..1293  
US-08-476-008-43

Query Match 1.6%; Score 32.8; DB 1; Length 1293;  
Best Local Similarity 45.9%; Pred. NO. 1.7;  
Matches 112; Conservative 0; Mismatches 132; Indels 0; Gaps 0;  
QY 370 ttcaaacccgttgatggcaaatgagacgacctatcactactactatgaggaataatgg 429  
DB 764 TTCATAATGTTGGAATCAATCAACACCGTTCAGGTATTATTGATTTGTTGAAAAAATGG 823  
QY 430 gatccactgaattgatcgatctcaatgggtgcgcatgggctcgctctttcttctgggt 489  
DB 824 GGGGTAAATCAACTTTTCAATCAACACACTGGTGTGCTGAACCTACTGCTTCTATTCGTA 883  
QY 490 cgcagctctaccgcgcgaactatcttgaacattctcaaacgcgtattcggaagaacat 549  
DB 884 TTCAATACACACCAATGCTTCAACCAATAACCAATGAAGGAGATTAGTTCCAAAAGCAA 943  
QY 550 tatggcaaatctgttaagaagctctcttctaccacagcgtttacaaagactacaaga 609  
DB 944 TTGATGAAGTGCCTGTAAATAGCATTACTTTGTACACAGCAGTGGCAGGTACAAATTA 1003  
QY 610 gaga 613  
DB 1004 AAGA 1007

RESULT 13  
US-08-306-063-43  
Sequence 43, Application US/08306063  
Patent No. 5633435  
GENERAL INFORMATION:  
APPLICANT: Barry, Gerard F.  
APPLICANT: Kishore, Ganesh M.  
APPLICANT: Padgett, Stephen R.  
APPLICANT: Stallings, William C.  
TITLE OF INVENTION: Glyphosate Tolerant  
NUMBER OF SEQUENCES: 69  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Dennis R. Hoerner, Jr., Monsanto Co. BB4F  
STREET: 700 Chesterfield Village Parkway  
CITY: St. Louis  
STATE: Missouri  
COUNTRY: USA  
ZIP: 63198  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/306,063  
FILING DATE: 13-SEP-1994  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/749,611  
FILING DATE: 28-AUG-1991  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Hoerner Jr., Dennis R.  
REGISTRATION NUMBER: 30,914  
REFERENCE/DOCKET NUMBER: 38-21(10660)A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (314)537-6099  
TELEFAX: (314)537-6047

```

; INFORMATION FOR SEQ ID NO: 43:
;
; SEQUENCE CHARACTERISTICS:
;     LENGTH: 1293 base pairs
;     TYPE: nucleic acid
;     STRANDEDNESS: double
;     TOPOLOGY: linear
;     MOLECULE TYPE: DNA (genomic)
;     FEATURE:
;         NAME/KEY: CDS
;         LOCATION: 1..1293
;
US-08-306-063-43

```

Query Match	1.6%	Score 32.8;	DB 1;	Length 1293;
Best Local Similarity	45.9%	Pred. No. 1.7;		
Matches 112;	Conservative 0;	Mismatches 132;	Indels 0;	Gaps 0;
QY	370	ttcacaaaccgttgatggcaaatgagacgcgctatcacctactactatgaggaaaactatgg	429	
Db)	764	TTCATATGTTGGAATCAATCAACACGTTTCAGGTATTATGATATTCTTCAAAAATGG	823	
QY	430	gattccacttgaaattgatgcgatctcaatgggtgcgcattggggcctctgctctctctctg	489	
Db	824	CGGTAATATCCAACTTTTTCATCAAAACAACGTGGTGTGAACCTACTGCTTCTATTCTGA	883	
QY	490	cggagtccttaccgcgcgaactatcttgaacattctcaaacgcgtatttcggcgaagaacat	549	
Db	884	TTCAATACACACCAATGCTTCAACCAATAACAATCGAAGGAGAAATTAGTTCCAAAAGCAA	943	
QY	550	tatggcaaaattctgtttaagaagtcctcttatcccaagaagctttacaaagaactacaa	609	
Db	944	TTGATGAACGTGCCTGTAAATAGCATTTGTACACAAAGCAGTTGGCACGAGTACAA	1003	
QY	610	gaga	613	
Db	1004	AAGA	1007	

RESULT 14  
US-08-833-485-43  
; Sequence 43, Application US/08833485  
; Patent No. 5804425  
; GENERAL INFORMATION:  
; APPLICANT: Barry, Gerard F.  
; APPLICANT: Kishore, Ganesh M.  
; APPLICANT: Padgett, Stephen R.  
; APPLICANT: Stallings, William C.  
; TITLE OF INVENTION: Glyphosate Tolerant  
; TITLE OF INVENTION: 5-Enolpyruvylshikimate-3-Phosphate Synthases  
; NUMBER OF SEQUENCES: 69  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Dennis R. Hoerner, Jr., Monsanto Co. BBAF  
; STREET: 700 Chesterfield Village Parkway  
; CITY: St. Louis  
; STATE: Missouri  
; COUNTRY: USA  
; ZIP: 63198  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/833,485  
; FILING DATE: 07-APR-1997  
; CLASSIFICATION: 435  
; PRIORITY APPLICATION DATA:  
; APPLICATION NUMBER: US 08/306,063  
; FILING DATE: 13-SEP-1994  
; CLASSIFICATION: 435  
; PRIORITY APPLICATION DATA:  
; APPLICATION NUMBER: US 07/749,611  
; FILING DATE: 28-AUG-1991

```

; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/576,537
; FILING DATE: 31-AUG-1990
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Hoerner Jr., Dennis R.
; REGISTRATION NUMBER: 30,914
; REFERENCE/DOCKET NUMBER: 38-21(15117)A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (314)737-6099
; TELEFAX: (314)737-6047
; INFORMATION FOR SEQ ID NO: 43:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1293 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1293
; US-08-833-485-43

Query Match 1.6%; Score 32.8; DB 1; Length 1293;
Best Local Similarity 45.9%; Pred. No. 1..7;
Matches 112; Conservative 0; Mismatches 132; Indels 0; Gaps 0;

QY 370 ttcaaacaccttgatggcgaaatgagacggcctatcactactatgatgaggaactatgg 429
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 764 TTCATAATGTTGGAAATCAATCAACACGTTTCAGGTATTATTATGATATTGTTGAAAAAATGG 823

QY 430 gatccacttgaattgatgcgattcctaattgctgcgcatggggcctcgctctcttcgggt 489
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 824 GCGGTAATATCCAACTTTTTCAAATCAAAACAACGTGCTGAACCTACTGCTTCATTTCGTA 883

QY 490 cgcagcttcctacgcgcgaactatcttgaacattctcaaacgggtattcggcagaagaacat 549
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 884 TTCATAACACACCAATGCTTCAACCAATACATCGAAGGAGAAATTAGTTCCTCAAAACCAA 943

QY 550 tatggcaaatcttgtaagaagctcttattctaccgaagcgtttacaagaactacaacga 609
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 944 TTGATGAAGCTGCCTGTATAGCATTACTTTGTACACAAAGCAGTTGGCAGCAGTACAATTA 1003

QY 610 gaga 613
|||
Db 1004 AAGA 1007

RESULT 15
US-09-137-440-43
; Sequence 43, Application US/09137440
; Patent No. 6248876
; GENERAL INFORMATION:
; APPLICANT: Barry, Gerard F.
; APPLICANT: Kishore, Ganesh M.
; APPLICANT: Padgett, Stephen R.
; APPLICANT: Stallings, William C.
; TITLE OF INVENTION: Glyphosate Tolerant
; TITLE OF INVENTION: 5-Enolpyruvylshikimate-3-Phosphate Synthases
; NUMBER OF SEQUENCES: 69
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dennis R. Hoerner, Jr., Monsanto Co. BB4F
; STREET: 700 Chesterfield Village Parkway
; CITY: St. Louis
; STATE: Missouri
; COUNTRY: USA
; ZIP: 63198
; COMPUTER READABLE FORM: disk
; MEDIUM TYPE: Floppy
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

```



```

1 SOFTWARE: PatentIn Release #1.0, Version #1.25
2 CURRENT APPLICATION DATA:
3 APPLICATION NUMBER: US/09/137,440
4 FILING DATE:
5 CLASSIFICATION:
6 PRIOR APPLICATION DATA:
7 APPLICATION NUMBER: US 08/833,485
8 FILING DATE: 07-APR-1997
9 APPLICATION NUMBER: US 08/306,063
10 FILING DATE: 13-SEP-1994
11 PRIOR APPLICATION DATA:
12 APPLICATION NUMBER: US 07/749,611
13 FILING DATE: 28-AUG-1991
14 PRIOR APPLICATION DATA:
15 APPLICATION NUMBER: US 07/576,537
16 FILING DATE: 31-AUG-1990
17 ATTORNEY/AGENT INFORMATION:
18 NAME: Hoerner Jr., Dennis R.
19 REGISTRATION NUMBER: 30,914
20 REFERENCE/DOCKET NUMBER: 38-21(15117)A
21 TELEPHONE: (314)737-6099
22 TELEFAX: (314)737-6047
23 INFORMATION FOR SEQ ID NO: 43:
24 SEQUENCE CHARACTERISTICS:
25 LENGTH: 1293 base pairs
26 TYPE: nucleic acid
27 STRANDEDNESS: double
28 TOPOLOGY: linear
29 MOLECULE TYPE: DNA (genomic)
30 FEATURE:
31 NAME/KEY: CDS
32 LOCATION: 1..1293
33 US-09-137-440-43

```

Query Match	1.6%	Score 32.8;	DB 4;	Length 1293;
Best Local Similarity	45.9%;	Pred. No. 1.7;		
Matches 112;	Conservative 0;	Mismatches 132;	Indels 0;	Gaps 0;
Qy 370	ttceacaaccgttgatggcaaatgagacggcctatcactactactatgagggaaaactatgg	429		
Db 764	TTCAATAATGTTGGAAATCAATCAAAACACGTCAGGTATATTGATATATTGCGAAAAATGG	823		
Qy 430	gatccacttgaattgatgcgactcaatggtgcgcgatggggcctcgctcttctcttggt	489		
Db 824	CGGTAATATCCAACTTTTCAATCAAAACACTGGTGTGTAACCTACGCTCTTCTATTTCGTA	883		
Qy 490	cgcagctctaccgcgcgaactatcttgaacattctcaaacgcgtattcggcgaagaacat	549		
Db 884	TTCAATACACACCAATGCTTCAACCANTACAAATCGAAGGAGANTAGTTCCAAAGCAA	943		
Qy 550	tatggcaaatcttgttaagaagtctcttatctaccacagcgtttcacaaagactacaacga	609		
Db 944	TTGATGAACCTGCCGTGTAATAGCAATTACTTTGTACACAAGCAGTTGGCAGCAGTACAAATTA	1003		
Qy 610	gaga	613		
Db 1004	AAGA	1007		

Search completed: May 14, 2002, 07:25:16  
Job time: 10843 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 14, 2002, 05:52:23 ; Search time 186.08 Seconds  
(without alignments)  
9638.459 Million cell updates/sec

Title: US-09-646-825-1

Perfect score: 2092

Sequence: 1 gaattcttagactccacca.....cttggttaagagctcaagctt 2092

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 930621 seqs, 428662619 residues

Total number of hits satisfying chosen parameters: 1861242

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- N\_Geneseq\_1101.\*
- 1: /SIDSI/qcgdata/geneseq/geneseq/NA1980.DAT.\*
  - 2: /SIDSI/qcgdata/geneseq/geneseq/NA1981.DAT.\*
  - 3: /SIDSI/qcgdata/geneseq/geneseq/NA1982.DAT.\*
  - 4: /SIDSI/qcgdata/geneseq/geneseq/NA1983.DAT.\*
  - 5: /SIDSI/qcgdata/geneseq/geneseq/NA1984.DAT.\*
  - 6: /SIDSI/qcgdata/geneseq/geneseq/NA1985.DAT.\*
  - 7: /SIDSI/qcgdata/geneseq/geneseq/NA1986.DAT.\*
  - 8: /SIDSI/qcgdata/geneseq/geneseq/NA1987.DAT.\*
  - 9: /SIDSI/qcgdata/geneseq/geneseq/NA1988.DAT.\*
  - 10: /SIDSI/qcgdata/geneseq/geneseq/NA1989.DAT.\*
  - 11: /SIDSI/qcgdata/geneseq/geneseq/NA1990.DAT.\*
  - 12: /SIDSI/qcgdata/geneseq/geneseq/NA1991.DAT.\*
  - 13: /SIDSI/qcgdata/geneseq/geneseq/NA1992.DAT.\*
  - 14: /SIDSI/qcgdata/geneseq/geneseq/NA1993.DAT.\*
  - 15: /SIDSI/qcgdata/geneseq/geneseq/NA1994.DAT.\*
  - 16: /SIDSI/qcgdata/geneseq/geneseq/NA1995.DAT.\*
  - 17: /SIDSI/qcgdata/geneseq/geneseq/NA1996.DAT.\*
  - 18: /SIDSI/qcgdata/geneseq/geneseq/NA1997.DAT.\*
  - 19: /SIDSI/qcgdata/geneseq/geneseq/NA1998.DAT.\*
  - 20: /SIDSI/qcgdata/geneseq/geneseq/NA1999.DAT.\*
  - 21: /SIDSI/qcgdata/geneseq/geneseq/NA2000.DAT.\*
  - 22: /SIDSI/qcgdata/geneseq/geneseq/NA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	DB	ID	Description
1	2092	100.0	2092	20	AAZ59028		Codon optimised ye
2	1239.8	59.3	2059	20	AAZ59061		Wild type yeast Fe
3	83	4.0	83	20	AAZ59031		Oligonucleotide A-
4	83	4.0	83	20	AAZ59032		Oligonucleotide A-
5	83	4.0	83	20	AAZ59033		Oligonucleotide A-
6	83	4.0	83	20	AAZ59034		Oligonucleotide A-
7	83	4.0	83	20	AAZ59035		Oligonucleotide A-
8	83	4.0	83	20	AAZ59036		Oligonucleotide A-
9	83	4.0	83	20	AAZ59043		Oligonucleotide C-
10	83	4.0	83	20	AAZ59044		Oligonucleotide C-
11	83	4.0	83	20	AAZ59045		Oligonucleotide C-

C 12	83	4.0	83	20	AAZ59046	Oligonucleotide C-
C 13	83	4.0	83	20	AAZ59047	Oligonucleotide C-
C 14	83	4.0	83	20	AAZ59048	Oligonucleotide C-
C 15	82	3.9	82	20	AAZ59049	Oligonucleotide D-
C 16	82	3.9	82	20	AAZ59050	Oligonucleotide D-
C 17	82	3.9	82	20	AAZ59051	Oligonucleotide D-
C 18	82	3.9	82	20	AAZ59052	Oligonucleotide D-
C 19	82	3.9	82	20	AAZ59053	Oligonucleotide D-
C 20	82	3.9	82	20	AAZ59054	Oligonucleotide D-
C 21	80.6	3.9	936	22	AAF58252	Oligonucleotide D1
C 22	80.6	3.9	936	22	AAF58253	Oligonucleotide D1
C 23	80.6	3.9	936	22	AAF58257	Oligonucleotide D1
C 24	80.6	3.9	936	22	AAF58259	Oligonucleotide D2
C 25	80.6	3.9	936	22	AAF58262	Oligonucleotide D2
C 26	80.6	3.9	938	22	AAF58255	Oligonucleotide D1
C 27	80	3.8	80	20	AAZ59037	Oligonucleotide B-
C 28	80	3.8	80	20	AAZ59038	Oligonucleotide B-
C 29	80	3.8	80	20	AAZ59039	Oligonucleotide B-
C 30	80	3.8	80	20	AAZ59040	Oligonucleotide B-
C 31	80	3.8	80	20	AAZ59041	Oligonucleotide B-
C 32	80	3.8	80	20	AAZ59042	Oligonucleotide B-
C 33	77	3.7	77	20	AAZ59055	Oligonucleotide E-
C 34	77	3.7	77	20	AAZ59056	Oligonucleotide E-
C 35	77	3.7	77	20	AAZ59057	Oligonucleotide E-
C 36	77	3.7	77	20	AAZ59058	Oligonucleotide E-
C 37	77	3.7	77	20	AAZ59059	Oligonucleotide E-
C 38	77	3.7	77	20	AAZ59060	Oligonucleotide E-
C 39	75.6	3.6	936	22	AAF58252	Oligonucleotide D1
C 40	75.6	3.6	936	22	AAF58254	Oligonucleotide D1
C 41	75.6	3.6	936	22	AAF58257	Oligonucleotide D1
C 42	75.6	3.6	936	22	AAF58259	Oligonucleotide D2
C 43	75.6	3.6	936	22	AAF58262	Oligonucleotide D2
C 44	75.6	3.6	938	22	AAF58255	Oligonucleotide D1
C 45	43	2.1	4590	22	AAH24065	Yeast AOD9604-asso

ALIGNMENTS

RESULT 1

AAZ59028	AAZ59028 standard; DNA; 2092 BP.
ID	AAZ59028 standard; DNA; 2092 BP.
XX	
AC	AAZ59028;
XX	
DT	11-APR-2000 (first entry)
XX	
DE	Codon optimised yeast Fe3+-reductase gene.
XX	
KW	Crop plant; polyadenylation signal; substitution; ferric reductase; iron;
KW	yeast; heterologous gene; ds.
XX	
OS	Saccharomyces cerevisiae.
OS	Synthetic.
XX	
FH	Key
FT	CDS
FT	Location/Qualifiers
FT	20..2080
FT	/*tag= a
FT	/product= "Fe3+-reductase"
XX	
PN	WO9948356-A1.
XX	
PD	30-SEP-1999.
XX	
PF	24-MAR-1999; 99WO-JP01481.
XX	
PR	24-MAR-1998; 98JP-0096637.
XX	
PA	(NISC-) JAPAN SCI & TECHNOLOGY CORP.
XX	
PI	Mori S, Nakanishi H, Oki H, Yamaguchi H;
XX	
DR	WPI; 1999-571924/48.

DR P-PSDB; AAY56088.

XX	f-PSUB; RAHJ00086.
PT	Transformation of higher plant with foreign gene having modified poly(A) site -

PS Claim 20; Fig 9; 81pp; Japanese.

The invention relates to a method of transforming a higher plant with a foreign gene in which the polyadenylation signal region contained in the nucleotide sequence of the foreign gene has been modified by nucleotide substitution to give a sequence not having this function, in such a way that the function of the foreign gene expression product is not substantially impaired. The foreign gene used is especially an iron (Fe) 3+ (ferric) reductase gene in which putative polyadenylation signals contained within the coding sequence are substituted with alternative codons for expression in higher plants. This sequence represents a plant codon optimised version of the wild type yeast ferric reductase (FRR1 gene). The sequence has had putative polyadenylation signal and poly(A) addition sites removed by nucleotide substitution. The method allows the production of transformant plants with improved expression of a desired foreign gene, especially to improve iron absorption efficiency in crop plants.

Sequence 2092 BP: 587 A: 510 C: 418 G: 577 T: 0 other;

Query Match	100.0%;	Score 2092;	DB 20;	Length 2092;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 2092:	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

Qy	1	gaattctctagactccaacatggttagaacaacagagtcctttctgctcttcatctcttt	60
Db	1	gaattctctagactccaacatggttagaacaacagagtcctttctgctcttcatctcttt	60
Qy	61	cttcgtcacagtcacacatcgagcgtctacactcatctccaactctcatcttcagagctgc	120
Db	61	cttcgtcacagtcacacatcgagcgtctacactcatctccaactctcatcttcagagctgc	120
Qy	121	actgtacacagtcctggatgctcgaagcaagtcacaagcttctgtactgtcaagaacatcaattg	180
Db	121	actgtacacagtcctggatgctcgaagcaagtcacaagcttctgtactgtcaagaacatcaattg	180
Qy	181	gctcggaaagcgtcactgcactgccttatgagaactccaaatctcaacaaagctctgcactc	240
Db	181	gctcggaaagcgtcactgcactgccttatgagaactccaaatctcaacaaagctctgcactc	240
Qy	241	cgctttgatgaaccttgccagccaactgtcaagtatcaaggtttacacactggaggacat	300
Db	241	cgctttgatgaaccttgccagccaactgtcaagtatcaaggtttacacactggaggacat	300
Qy	301	gaagaacaactacacttaatgcaagtaactacccttcgctctcttgagaaatccgataagaa	360
Db	301	gaagaacaactacacttaatgcaagtaactacccttcgctctcttgagaaatccgataagaa	360
Qy	361	gacagtgtttccaaacgcttgatggcaaatgagcggcctatcactactactatgagga	420
Db	361	gacagtgtttccaaacgcttgatggcaaatgagcggcctatcactactactatgagga	420
Qy	421	aaactatgggataccacttgaaatttgaatcgatctcaatggtcgcatggggcctctctt	480
Db	421	aaactatgggataccacttgaaatttgaatcgatctcaatggtcgcatggggcctctctt	480
Qy	481	cttctgggtcgcagtccttaccgcgcgaactatcttgaaactctctcaaacgcgtattcgg	540
Db	481	cttctgggtcgcagtccttaccgcgcgaactatcttgaaactctctcaaacgcgtattcgg	540
Qy	541	caagaacatattgcaaaattctgttaagaagctctcttatcccccaagcgtttacaaaga	600
Db	541	caagaacatattgcaaaattctgttaagaagctctcttatcccccaagcgtttacaaaga	600
Qy	601	ctacaacagagaaactttctatcttttggaaacggttgccattcaactttacaactcgagg	660
Db	601	ctacaacagagaaactttctatcttttggaaacggttgccattcaactttacaactcgagg	660





ID AA259032 standard; DNA; 83 BP.  
 XX AC AA259032;  
 XX DT 11-APR-2000 (first entry)  
 XX DE Oligonucleotide A-2 for generating codon-optimised ferric reductase.  
 XX Crop plant; polyadenylation signal; substitution; ferric reductase; iron;  
 KW yeast; heterologous gene; ss.  
 XX OS Saccharomyces cerevisiae.  
 OS Synthetic.  
 XX PN WO9948356-A1.  
 XX PD 30-SEP-1999.  
 XX PF 24-MAR-1999; 99WO-JP01481.  
 XX PR 24-MAR-1998; 98JP-0096637.  
 XX PA (NISC-) JAPAN SCI & TECHNOLOGY CORP.  
 XX PI Mori S, Nakanishi H, Oki H, Yamaguchi H;  
 XX WPI; 1999-571924/48.  
 XX PT Transformation of higher plant with foreign gene having modified  
 PT poly(A) site -  
 XX PS Disclosure; Fig 6; 81pp; Japanese.  
 XX CC The invention relates to a method of transforming a higher plant with  
 CC a foreign gene in which the polyadenylation signal region contained in  
 CC the nucleotide sequence of the foreign gene has been modified by  
 CC nucleotide substitution to give a sequence not having this function, in  
 CC such a way that the function of the foreign gene expression product is  
 CC not substantially impaired. The foreign gene used is especially an iron  
 CC (Fe) 3+ (ferric) reductase gene in which putative polyadenylation  
 CC signals contained within the coding sequence are substituted with  
 CC alternative codons for expression in higher plants. Oligonucleotides  
 CC AA259031-259060 were used as overlapping oligonucleotide primers to  
 CC generate a plant codon-optimised yeast ferric reductase (FRE1 gene;  
 CC AA259028) in which putative polyadenylation signal and poly(A) addition  
 CC sites are removed by nucleotide substitution. The method allows the  
 CC production of transformant plants with improved expression of a desired  
 CC foreign gene, especially to improve iron absorption efficiency in crop  
 CC plants.  
 XX SQ Sequence 83 BP; 21 A; 26 C; 15 G; 21 T; 0 other;  
 Query Match 4.0%; Score 83; DB 20; Length 83;  
 Best Local Similarity 100.0%; Pred. No. 5.2e-16;  
 Matches 83; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 71 gtccaatcagcgctacacatcctcactcgcattcgcatttcaggctgacctgaccag 130  
 Db 1 gtccaatcagcgctacacatcctcactcgcattcgcatttcaggctgacctgaccag 60  
 QY 131 ttccggtgctcgaagcagtcacaa 153  
 Db 61 ttccggtgctcgaagcagtcacaa 83  
 RESULT 5  
 ID AA259033  
 XX AA259033 standard; DNA; 83 BP.  
 XX AC AA259033;  
 XX DT 11-APR-2000 (first entry)

XX Oligonucleotide A-3 for generating codon-optimised ferric reductase.  
 XX Crop plant; polyadenylation signal; substitution; ferric reductase; iron;  
 KW yeast; heterologous gene; ss.  
 XX OS Saccharomyces cerevisiae.  
 OS Synthetic.  
 XX PN WO9948356-A1.  
 XX PD 30-SEP-1999.  
 XX PF 24-MAR-1999; 99WO-JP01481.  
 XX PR 24-MAR-1998; 98JP-0096637.  
 XX PA (NISC-) JAPAN SCI & TECHNOLOGY CORP.  
 XX PI Mori S, Nakanishi H, Oki H, Yamaguchi H;  
 XX WPI; 1999-571924/48.  
 XX PT Transformation of higher plant with foreign gene having modified  
 PT poly(A) site -  
 XX PS Disclosure; Fig 6; 81pp; Japanese.  
 XX CC The invention relates to a method of transforming a higher plant with  
 CC a foreign gene in which the polyadenylation signal region contained in  
 CC the nucleotide sequence of the foreign gene has been modified by  
 CC nucleotide substitution to give a sequence not having this function, in  
 CC such a way that the function of the foreign gene expression product is  
 CC not substantially impaired. The foreign gene used is especially an iron  
 CC (Fe) 3+ (ferric) reductase gene in which putative polyadenylation  
 CC signals contained within the coding sequence are substituted with  
 CC alternative codons for expression in higher plants. Oligonucleotides  
 CC AA259031-259060 were used as overlapping oligonucleotide primers to  
 CC generate a plant codon-optimised yeast ferric reductase (FRE1 gene;  
 CC AA259028) in which putative polyadenylation signal and poly(A) addition  
 CC sites are removed by nucleotide substitution. The method allows the  
 CC production of transformant plants with improved expression of a desired  
 CC foreign gene, especially to improve iron absorption efficiency in crop  
 CC plants.  
 XX SQ Sequence 83 BP; 26 A; 21 C; 17 G; 19 T; 0 other;  
 Query Match 4.0%; Score 83; DB 20; Length 83;  
 Best Local Similarity 100.0%; Pred. No. 5.2e-16;  
 Matches 83; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 141 caagcaagtcacagcttctgctactgcaagaacatcattggtcgcgaagcgctcactgcat 200  
 Db 1 caagcaagtcacagcttctgctactgcaagaacatcattggtcgcgaagcgctcactgcat 60  
 QY 201 gcgcttatgagaactcacaatct 223  
 Db 61 gcgcttatgagaactcacaatct 83  
 RESULT 6  
 ID AA259034/C  
 XX AA259034 standard; DNA; 83 BP.  
 XX AC AA259034;  
 XX DT 11-APR-2000 (first entry)  
 XX DE Oligonucleotide A-4 for generating codon-optimised ferric reductase.  
 XX Crop plant; polyadenylation signal; substitution; ferric reductase; iron;  
 KW yeast; heterologous gene; ss.







XX Disclosure; Fig 6; 8lpp; Japanese.

XX The invention relates to a method of transforming a higher plant with

CC a foreign gene in which the polyadenylation signal region contained in

CC the nucleotide sequence of the foreign gene has been modified by

CC nucleotide substitution to give a sequence not having this function, in

CC such a way that the function of the foreign gene expression product is

CC not substantially impaired. The foreign gene used is especially an iron

CC (Fe) 3+ (ferric) reductase gene in which putative polyadenylation

CC signals contained within the coding sequence are substituted with

CC alternative codons for expression in higher plants. Oligonucleotides

CC AAZ59031-259060 were used as overlapping oligonucleotide primers to

CC generate a plant codon-optimised yeast ferric reductase (FRE1 gene;

CC AAZ59028) in which putative polyadenylation signal and poly(A) addition

CC sites are removed by nucleotide substitution. The method allows the

CC production of transformant plants with improved expression of a desired

CC foreign gene, especially to improve iron absorption efficiency in crop

CC plants.

Sequence 83 BP; 18 A; 21 C; 18 G; 26 T; 0 other;

Query Match 4.0%; Score 83; DB 20; Length 83;

Best Local Similarity 100.0%; Pred. No. 5.2e-16;

Matches 83; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 911 gcatacgtcgttcattgtagcgtgcgttcattcaatcgttatgacgcgttcaggagtt 970

DB 1 gcatacgtcgttcattgtagcgtgcgttcattcaatcgttatgacgcgttcaggagtt 60

QY 971 aaacgagagattccagtcctc 993

DB 61 aaacgagagattccagtcctc 83

RESULT 11

AAZ59045

ID AAZ59045 standard; DNA; 83 BP.

XX

AC AAZ59045;

XX

XX 11-APR-2000 (first entry)

XX Oligonucleotide C-4 for generating codon-optimised ferric reductase.

DE Crop plant; polyadenylation signal; substitution; ferric reductase; iron;

XX yeast; heterologous gene; ss.

OS Saccharomyces cerevisiae.

XX Synthetic.

PN WO9948356-A1.

XX

XX 30-SEP-1999.

XX

XX 24-MAR-1999; 99WO-JP01481.

XX

XX 24-MAR-1998; 98JP-0096637.

XX

XX (NISC-) JAPAN SCI & TECHNOLOGY CORP.

XX

XX Mori S, Nakanishi H, Oki H, Yamaguchi H;

XX WPI; 1999-571924/48.

XX

XX Transformation of higher plant with foreign gene having modified

PT poly(A) site -

XX

XX Disclosure; Fig 6; 8lpp; Japanese.

PS

XX The invention relates to a method of transforming a higher plant with

CC a foreign gene in which the polyadenylation signal region contained in

CC the nucleotide sequence of the foreign gene has been modified by

CC nucleotide substitution to give a sequence not having this function, in

CC such a way that the function of the foreign gene expression product is

CC not substantially impaired. The foreign gene used is especially an iron

CC (Fe) 3+ (ferric) reductase gene in which putative polyadenylation

CC signals contained within the coding sequence are substituted with

CC alternative codons for expression in higher plants. Oligonucleotides

CC AAZ59031-259060 were used as overlapping oligonucleotide primers to

CC generate a plant codon-optimised yeast ferric reductase (FRE1 gene;

CC AAZ59028) in which putative polyadenylation signal and poly(A) addition

CC sites are removed by nucleotide substitution. The method allows the

CC production of transformant plants with improved expression of a desired

CC foreign gene, especially to improve iron absorption efficiency in crop

CC plants.

XX

SQ Sequence 83 BP; 22 A; 19 C; 13 G; 29 T; 0 other;

Query Match 4.0%; Score 83; DB 20; Length 83;

Best Local Similarity 100.0%; Pred. No. 5.2e-16;

Matches 83; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 981 tattccagtcctctgtaaggaaattctacttcagatggggaatagtagccacaattctta 1040

DB 1 tattccagtcctctgtaaggaaattctacttcagatggggaatagtagccacaattctta 60

QY 1041 tgtccatcatcatttccagtc 1063

DB 61 tgtccatcatcatttccagtc 83

RESULT 12

AAZ59046/C

ID AAZ59046 standard; DNA; 83 BP.

XX

AC AAZ59046;

XX

XX 11-APR-2000 (first entry)

XX Oligonucleotide C-4 for generating codon-optimised ferric reductase.

DE Crop plant; polyadenylation signal; substitution; ferric reductase; iron;

XX yeast; heterologous gene; ss.

OS Saccharomyces cerevisiae.

XX Synthetic.

PN WO9948356-A1.

XX

XX 30-SEP-1999.

XX

XX 24-MAR-1999; 99WO-JP01481.

XX

XX 24-MAR-1998; 98JP-0096637.

XX

XX (NISC-) JAPAN SCI & TECHNOLOGY CORP.

XX

XX Mori S, Nakanishi H, Oki H, Yamaguchi H;

XX WPI; 1999-571924/48.

XX

XX Transformation of higher plant with foreign gene having modified

PT poly(A) site -

XX

XX Disclosure; Fig 6; 8lpp; Japanese.

PS

XX The invention relates to a method of transforming a higher plant with

CC a foreign gene in which the polyadenylation signal region contained in

CC the nucleotide sequence of the foreign gene has been modified by

CC nucleotide substitution to give a sequence not having this function, in

CC such a way that the function of the foreign gene expression product is

CC not substantially impaired. The foreign gene used is especially an iron

CC (Fe) 3+ (ferric) reductase gene in which putative polyadenylation



SQ Sequence 83 BP; 23 A; 20 C; 23 G; 17 T; 0 other;

Query Match 4.0%; Score 83; DB 20; Length 83;  
Best Local Similarity 100.0%; Pred. No. 5.2e-16;  
Matches 83; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1193 gctggcattctctgttcgacaggttctgcgaattgtacgtatcatcatgaacggagt 1252  
|||||  
Db 83 CTTGATCCTCTGCTTCGACAGGTTCTGCCGATTGTACGTATCATCATGAACGGAGGT 24

Oy 1253 cttagaccgcgaactttgttcgac 1275  
|||||  
Db 23 CTTAAGACCGCCACTTTGTGCAC 1

RESULT 15  
AAZ59049  
ID AAZ59049 standard; DNA; 82 BP.

AAZ59049;

DT 11-APR-2000 (first entry)

XX Oligonucleotide D-1 for generating codon-optimised ferric reductase.

DE Crop plant; polyadenylation signal; substitution; ferric reductase; iron;  
KW yeast; heterologous gene; ss.

XX Saccharomyces cerevisiae.

OS Synthetic.

XX WO9948356-A1.

XX 30-SEP-1999.

XX 24-MAR-1999; 99WO-JP01481.

XX 24-MAR-1998; 98JP-0096637.

XX (NISC-) JAPAN SCI & TECHNOLOGY CORP.

XX Mori S, Nakanishi H, Oki H, Yamaguchi H;

XX WPI; 1999-571924/48.

XX Transformation of higher plant with foreign gene having modified

PT poly(A) site -

Disclosure; Fig 6; 81pp; Japanese.

XX The invention relates to a method of transforming a higher plant with  
CC a foreign gene in which the polyadenylation signal region contained in  
CC the nucleotide sequence of the foreign gene has been modified by  
CC nucleotide substitution to give a sequence not having this function, in  
CC such a way that the function of the foreign gene expression product is  
CC not substantially impaired. The foreign gene used is especially an iron  
CC (Fe) 3+ (ferric) reductase gene in which putative polyadenylation  
CC signals contained within the coding sequence are substituted with  
CC alternative codons for expression in higher plants. Oligonucleotides  
CC AAZ59031-259060 were used as overlapping oligonucleotide primers to  
CC generate a plant codon-optimised yeast ferric reductase (FRE1 gene;  
CC AAZ59028) in which putative polyadenylation signal and poly(A) addition  
CC sites are removed by nucleotide substitution. The method allows the  
CC production of transformant plants with improved expression of a desired  
CC foreign gene, especially to improve iron absorption efficiency in crop  
CC plants.

SQ Sequence 82 BP; 24 A; 18 C; 17 G; 23 T; 0 other;

Query Match 3.9%; Score 82; DB 20; Length 82;  
Best Local Similarity 100.0%; Pred. No. 1.1e-15;

us . . .

Matches 82; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Oy 1270 gtcgaccacacagatgattcttaacgtttatcaagatctctgtcaagaagcctaaagtcttcaa 1329  
|||||  
Db 1 gtcgaccacacagatgattcttaacgtttatcaagatctctgtcaagaagcctaaagtcttcaa 60  
Oy 1330 gtatcaagtgggagcatttgcc 1351  
|||||  
Db 61 gtatcaagtgggagcatttgcc 82

Search completed: May 14, 2002, 07:27:53  
Job time: 5730 sec

---



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 14, 2002, 04:19:53 ; Search time 2978.83 seconds  
(without alignments)  
11585.790 Million cell updates/sec

Title: US-09-646-825-1  
Perfect score: 2092  
Sequence: 1 gaattctctagactccacca.....cttggtaagagctcaagctt 2092

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0  
Searched: 1472140 seqs, 8248589755 residues  
Total number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl :  
1: gb\_ba :  
2: gb\_htg :  
3: gb\_in :  
4: gb\_ov :  
5: gb\_ov :  
6: gb\_pat :  
7: gb\_ph :  
8: gb\_pl :  
9: gb\_pr :  
10: gb\_ro :  
11: gb\_sts :  
12: gb\_sy :  
13: gb\_un :  
14: gb\_vi :  
15: em\_ba :  
16: em\_fun :  
17: em\_in :  
18: em\_in :  
19: em\_in :  
20: em\_or :  
21: em\_ov :  
22: em\_pat :  
23: em\_ph :  
24: em\_pl :  
25: em\_ro :  
26: em\_sts :  
27: em\_sy :  
28: em\_un :  
29: em\_vi :  
30: em\_htgo\_hum :  
31: em\_htgo\_inv :  
32: em\_htgo\_rod :  
33: em\_htgo\_rod :  
34: em\_htg\_hum :  
35: em\_htg\_inv :  
36: em\_htg\_rod :  
36: em\_htg\_other :  
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	2092	100.0	2092	6	E31006	E31006 Method for
2	1244	59.5	3264	8	YSCFRE1A	M86908 S. cerevisia
3	1244	59.5	38868	8	YSCHE167	U14913 Saccharomyc
4	147.6	7.1	694	6	A86082	A86082 Sequence 74
5	147.6	7.1	694	6	ARI55575	ARI55575 Sequence
6	147.6	7.1	694	6	E66100	E66100 Genome DNA
7	99.4	4.8	3084	8	D83181	D83181 Candida alb
8	97.8	4.7	3615	8	CAL387722	AJ387722 Candida a
9	90.2	4.3	711	6	A85974	A85974 Sequence 63
10	90.2	4.3	711	6	ARI55467	ARI55467 Sequence
11	90.2	4.3	711	6	E65992	E65992 Genome DNA
12	83	4.0	83	6	E31009	E31009 Method for
13	83	4.0	83	6	E31010	E31010 Method for
14	83	4.0	83	6	E31011	E31011 Method for
15	83	4.0	83	6	E31012	E31012 Method for
16	83	4.0	83	6	E31013	E31013 Method for
17	83	4.0	83	6	E31014	E31014 Method for
18	83	4.0	83	6	E31021	E31021 Method for
19	83	4.0	83	6	E31022	E31022 Method for
20	83	4.0	83	6	E31023	E31023 Method for
21	83	4.0	83	6	E31024	E31024 Method for
22	83	4.0	83	6	E31025	E31025 Method for
23	83	4.0	83	6	E31026	E31026 Method for
24	82	3.9	82	6	E31027	E31027 Method for
25	82	3.9	82	6	E31028	E31028 Method for
26	82	3.9	82	6	E31029	E31029 Method for
27	82	3.9	82	6	E31030	E31030 Method for
28	82	3.9	82	6	E31031	E31031 Method for
29	82	3.9	82	6	E31032	E31032 Method for
30	80	3.8	80	6	E31015	E31015 Method for
31	80	3.8	80	6	E31016	E31016 Method for
32	80	3.8	80	6	E31017	E31017 Method for
33	80	3.8	80	6	E31018	E31018 Method for
34	80	3.8	80	6	E31019	E31019 Method for
35	80	3.8	80	6	E31020	E31020 Method for
36	77	3.7	77	6	E31033	E31033 Method for
37	77	3.7	77	6	E31034	E31034 Method for
38	77	3.7	77	6	E31035	E31035 Method for
39	77	3.7	77	6	E31036	E31036 Method for
40	77	3.7	77	6	E31037	E31037 Method for
41	77	3.7	77	6	E31038	E31038 Method for
42	63	3.0	3987	8	SCYOR384W	Z75292 S.cerevisia
43	58.8	2.8	6955	8	SCYRL220C	Z28220 S.cerevisia
44	58.8	2.8	13213	8	SCSORE	X75950 S.cerevisia
45	53	2.5	974	11	CNS061Z7	AL400937 T7 end of

ALIGNMENTS

RESULT	1
E31006	E31006
LOCUS	2092 bp DNA
DEFINITION	Method for transforming plant, the thus transformed plant and gene thereof.
ACCESSION	E31006
VERSION	E31006.1 GI:13021492
KEYWORDS	JP 1999266876-A/1.
SOURCE	unidentified.
ORGANISM	unclassified.
REFERENCE	1 (bases 1 to 2092)
AUTHORS	Satoshi, M.H.O.O. and Nakanishi, H.Y.
TITLE	Method for transforming plant, the thus transformed plant and gene thereof
JOURNAL	Patent: JP 1999266876-A 1 05-OCT-1999;
COMMENT	SCIENCE & TECH AGENCY
	OS Unidentified
	PN JP 1999266876-A/1
	PD 05-OCT-1999





Db 1801 TGGGTTCATCAGTGGAGGATACAACTCAGATGAGTCCACTAGGGTTCGATGACAAGGA 1860  
 QY 1861 aqaatctgaatcacgttagaatgccttaacaagagccagacacctcaagagctagtcaq 1920  
 Db 1861 AGAATCTGAATCACCCTGATGATGCTTTAACAAGAGCGACACCTCAAGAGCTAGTGAG 1920  
 QY 1921 atcagagatcaaatgtcagaactcgagaactcgagaacacacatcactcttactactcgcggacc 1980  
 Db 1921 ATCAGAGATCAAAATGTCAGAACTCGAGAACAACAACATCACTTCTACTCATGCGGACC 1980  
 QY 1981 aqcgactttcaatgcagcacttttagaagtcagtttagaagagctatcgatttagctctgaa 2040  
 Db 1981 AGCGACTTTCATATGACGACTTTAGAAATGCGAGTTTGTACAAGGATATCGAATCTAGTCTGAA 2040  
 QY 2041 gatagatgcgaactagagagagagtttactcttggttaagagctcaagctt 2092  
 Db 2041 GATAGATGTCGAACCTAGAGGAGGAGATTTTACTTGGTAAGAGCTCAAGCTT 2092

RESULT 2  
 YSCFREIA 3264 bp DNA PLN 27-APR-1993  
 S. cerevisiae ferric reductase gene, complete cds.  
 M86908  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Saccharomyces cerevisiae DNA.  
 Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
 1 (bases 1 to 3264)  
 Dancis,A., Roman,D.G., Anderson,G.J., Hinnebusch,A.G. and  
 Klausner,R.D.  
 Ferric reductase of saccharomyces cerevisiae: Molecular  
 characterization, role in iron uptake, and transcriptional control  
 by iron  
 Proc. Natl. Acad. Sci. U.S.A. 89, 3869-3873 (1992)  
 92237270  
 FEATURES  
 source  
 1..3264  
 /organism="Saccharomyces cerevisiae"  
 /db\_xref="taxon:4932"  
 1..977  
 /gene="FRE1"  
 /note="confers iron regulated activity to heterologous  
 coding region"  
 1..3264  
 /gene="FRE1"  
 929..3264  
 /gene="FRE1"  
 /note="5' end mapped by primer extension, 3' end estimated  
 from mRNA size"  
 978..3038  
 /gene="FRE1"  
 /note="essential for ferric reductase activity"  
 /codon\_start=1  
 /product="ferric reductase"  
 /protein\_id="AAA34608.1"  
 /db\_xref="GI:171521"

translation="MTRVRLVLCFLISFATVQSSATLISCSIAALYQGCSSKS  
 KSCYKNLWGLSVTACAYENSKNTLDSALMKLASQSSIKVYTLDDMKNIYLNAS  
 NYLRAPEKSDKRTVVSQPLMANETAYHYIYEENYGIHLNLMRSOWCAGLVFVWAVL  
 TAATLNLKRVGKINMANSVKSLIYPSVKDYNERFVLMKLPENFTRCKGLV  
 VLIFVILILSLFSGHNKILPHVDRPRWRMSAFVSRADLMALFPPVILFGRIN  
 NPFIPTGLSESTFNFYHKSAYVCFMLAVVHSIVNTASGVRGVQSLVRKFFYRWG  
 IVATLTIISIIIFQSKVFNRRGYEIFLIHKNMIMFIAMYYHCITLGMWGIWSMA  
 GLCDFRCSRPVRIIMNGLKTATLSTDDSNVIKISKPKFKFYQVGFAYMYFLS  
 PKSAFYSFQSPHTVFLSERHPDNNPDLTMVYKANGITRVLISKVLSPNHTVDC  
 KIFLREPGYVTPHTAKLKNLVGAAGLVAAIYPHEVECLRPSTDLQHKFYIV  
 NDLSHLKWFENELWLKESCEVSIYTGSSVEDNSDESTKGFDDKKESEITVECLN  
 KRPDLKELVRSEIKLSENNNITFYSCGPATFNDFRNAVVGQIDSLKIDVELEEE  
 SFTW"

971 a 591 c 630 g 1072 t

BASE COUNT

ORIGIN  
 Query Match 59.5%; Score 1244; DB 8; Length 3264;  
 Best Local Similarity 75.0%; Pred. No. 0;  
 Matches 1556; Conservative 0; Mismatches 520; Indels 0; Gaps 0;  
 QY 16 caccatggttagaaccagagctcttctgcctcttctcattcttcttcgtcacagctca 75  
 Db 974 CAATATGTTAGAACCGGTGATTTATCTGCTTATTATATCTTTTGTGCTACGGTTCA 1033  
 QY 76 atcagagctcacatctcattcattcattcattcattcattcattcattcattcattc 135  
 Db 1034 ATCGAGTGTACACTTATTAGCACTTCATGTATTTCCCAAGCTGCGCTATACCAATTTGG 1093  
 QY 136 atgctcaagcaagtcataagctcttgcctactgcaagacatcaattgctcggagcgctac 195  
 Db 1094 ATGTTCTAGTAAATCTAAAGTTGCTACTGTATAAACAATCAATTTGGCTGGGTTCAAGTAC 1153  
 QY 196 tgcctgctcttatgagaactccaaatcctaaagactctggactcgcgtcttggatgaact 255  
 Db 1154 AGCATGTGCTATGAGATTCCTCAATCTAACAACACATAGACAGGCGCTTAAATGAAGTT 1213  
 QY 256 tgcagagcaatgctcaagtatcaaggttttacacactggaggacatgaagaacatctacat 315  
 Db 1214 AGCATCCCAATGTTCAAGCATCAAGTTTATCTTTAGAGGACATGAAGAATATTATTTT 1273  
 QY 316 taatcgaagtaactacctcgcctcctcgagaaatccgataagaagacagtggtttccaca 375  
 Db 1274 AAATCGCTCAAAATTTATTTGAGAGCACCTCGAGAAAAGTGTAAAAAACCGTGTAGTCA 1333  
 QY 376 accgttgatggcaaatgagagcgctctactactactactactactactactactactactg 435  
 Db 1334 ACCGCTCATGGGCAAGACAGACAGCTATCTATTATTATGAGGAAAATATTGTTGTTATCCA 1393  
 QY 436 cttgaatttgatgcgatctcaatggtgcgcattggggcctcgtctctctctctctctctc 495  
 Db 1394 TCTTAACCTAATGCGCTCTCAATGTTGGTCTGCTCTCTCTCTCTCTCTCTCTCTCTCT 1453  
 QY 496 cctaccgcgcgaactatctgaacattctcaaacgcgtattcgcgcaagaacattatggc 555  
 Db 1454 GCTTATCGACGACCTATCTTGAACATCTGAAAAGGGTGTGTTGGTAAGAATCATCATGCG 1513  
 QY 556 aaatctgttaagaagctctcttatcaccacgcgtttacaagactacaacagagagaaac 615  
 Db 1514 AAATCCGCTCAAAATCACTTATTATCTTCTGTTTACAAGATTATTAATGAAGCAAC 1573  
 QY 616 tttctatcttggaaacgcttggccattcactcactcactcactcactcactcactcactcact 675  
 Db 1574 TTTTATTTATGGAAGCGCTCTACCATTTAATTTTACAACCTCGAGGCAAGGGTCTCGCTGT 1633  
 QY 676 tcttattcttgcattctgcattctctcactctctcactctctcactcactcactcactcact 735  
 Db 1634 ATTAATTTTGTATTTTGACTATATATTCTCTCAGTTTGGTCAATAATATTAAACTTCC 1693  
 QY 736 acatcctacatagacctagatgagagaatcaatggcattcgtctcgcgcgtgctga 795  
 Db 1694 ACACCATATATAGCCCGAGATGGAGAAGATAGGCGCTTTGTGAGTCTGTAGAGCAGA 1753  
 QY 796 cttgatggcaatcgctcttctcccggtggtgacatttccgttatcccggaacaacccctt 855  
 Db 1754 CTTGATGGCCATTGGCACTTTTCCAGTAGCTCTATCTATTTCGGAATAAGAAATAATCCCTT 1813  
 QY 856 catccaatcaccgattgagcttttagttagttagttagttagttagttagttagttagttag 915  
 Db 1814 CATCCCTATAACAGGGCTTCTCTTTTCTACATTTAATTTATCAATAATATGTTGCTGCTA 1873  
 QY 916 cgtctgcttcattgagcgtcgctcattcattcattcattcattcattcattcattcattcatt 975  
 Db 1874 CGTTGTTTTCATGTTGGCGCTTGTACACTCAATTTGATGACCGCTCGGGAGTGAAG 1933  
 QY 976 aggagattcccgactctctgttaaggaaattctactcagatggggaatagtagcccaat 1035

Db 1934 AGGTGCTGTTTCAAAGTCTGGTTAGGAAATTTTACTTTTAGTGGGTATAGTGGCAACGAT 1993  
Qy 1036 tcttatgtccatcatcatttccagtcgagaaagtccttcaggaaacgaggttatgaaat 1095  
Db 1994 ATTAAATGCTATATATTTTCCAAAGTGAAGAAAGTATTTAGAAATAGAGGGTATGAGAT 2053  
Qy 1096 cttcttactattcacaaagccatgaacatcatctttatcatagctatgtattaccattg 1155  
Db 2054 ATTCTCTCTTATTCATAAAGCGATGAATATATGTTTCATTATGCCATGTACTACCATTTG 2113  
Qy 1156 ccacacactagtaggagtgctggatctgtgctccatggctggatctctctctcttcagacg 1215  
Db 2114 TCACACCTGGGCTGGATGGTGGATTTGGCTCAATGGCTGCTATTTTATGCTTTGATAGT 2173  
Qy 1216 gttctgcgaattatcagatcatcatcagaacgaggtctttaagaccgaccttttgtcgac 1275  
Db 2174 ATTCTGCAGGATTTGTAGAATAATCATGAATATGTTGGCTTGAAGAACTGCTACTTTGAGTAC 2233  
Qy 1276 cacagatgattctcaacggttatcaagatctctgtcaagaagcctaagttcttccaagtatca 1335  
Db 2234 CACTCATGATTCTAATGTTATTAATTTTCAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGT 2293  
Qy 1336 agtggagcatttgcctatagctctctctctctctcccaaaatcaagcctggttctcacagttt 1395  
Db 2294 AGTAGGAGCTTTCGCATACATGATTTCTTTATCACCAAAAGTGGTGTCTTATAGTTT 2353  
Qy 1396 tcaatctcatcctccacagtcctcatcagaagcagacagatctcacaacccagatca 1455  
Db 2354 CCATCATCATCTTACAGTATTTATCGGAACGACACCGCTGATCCAAACATCCAGATCA 2413  
Qy 1456 actaactatgacgtcaacagctcaacgggcatcagagagactctctttagcaaaagttct 1515  
Db 2414 ATTGACGATGACGTAAAGGCAATAAAGGTATCACTCGAGTGTGTTTATCGAAAGTTCT 2473  
Qy 1516 aagcgctccaaacataacacgttgatgcaagattttcttagaggacataatgctgtaac 1575  
Db 2474 AAGTGCTCCAATCATACTGTTGATGTTGAAATATTCCTTGAAGGCCCATATGTTGTAAC 2533  
Qy 1576 tgcctccatctgcaaaccttaagagaaatctagtagagtagctgcggccctcgccgt 1635  
Db 2534 GGTTCACATATCGCTAAGCTAAAGAAATCTGTTAGTGTAGCGCTGTTGGGTGT 2593  
Qy 1636 ggcagcctatcccccctttcgttagatgcttagatgcttagctgacactgacatcaactgca 1695  
Db 2594 TCGGCTATTTATCGCCTATTTGTCGAATGTTACGGTGTACCATCTACTGATCAACTTCA 2653  
Qy 1696 gcaacagttctactgagtcgcaacgaccttagtcaaccttaagtggttcgaaacgagct 1755  
Db 2654 GCATAAATTTTACTGGATTTGTAATGACCTATCCCATTTGAAATGGTTGAAATGAAAT 2713  
Qy 1756 acaatggccttaaggagaaatcttgtgaagtcctgtcatctcacactgggtcatcagtga 1815  
Db 2714 GCATGGTTAAGGAGAAAGTTGTGAAGTCTCAGTCATATATCTGTTCCAGTGTGTA 2773  
Qy 1816 ggaataaacctcagatgagtcactaaggggttcgatgacagaagaatactgaaatcac 1875  
Db 2774 GGACACAAATTCAGATGAGATGACAAAGGTTTGTATGATGATAAGAAAGAAACCGAAATCAC 2833  
Qy 1876 cgtagaatgccttaacagaagccagacctcaagagagctagtagatcagagatcaaat 1935  
Db 2834 TGTGTAATGCTCAATAAAGACCTGATTTTGAAGAAGTACTAGTGGCTCGGAAATAAAACT 2893  
Qy 1936 gtcagaactcagacaacaacatcacttctactcatcgccgaccagcagcttcaatga 1995  
Db 2894 CTCGAACCTAGAGAAATAATATATACCTTTTATCTCGCGGGCAGCAACGTTTAAACGA 2953  
Qy 1996 cgaatttaggaatgcagttgtacaaggtatcgattcttagtctgaagatagatgctgaact 2055  
Db 2954 CGATTTTGAATAATGCAGTGGTCCAAGGTATAGACTCTTCTCTTGAAGATGACGTTGAACT 3013  
Qy 2056 agagagagagagatttacttggttaagagctcaagct 2091  
Db 3014 AGAAGAGAAAGTTTACATGGTAAGGCCCTCTGTT 3049

## RESULT 3

YSC8167

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

YSC8167 38868 bp DNA PLN 22-AUG-1997  
Saccharomyces cerevisiae chromosome XII cosmid 8167.  
U14913 Y13138  
U14913.1 GI:544497

baker's yeast strain-S288C (AB972).

Saccharomyces cerevisiae

Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
Saccharomycetales; Saccharomycetaceae; Saccharomycetes.

1 (bases 1 to 38868)

Johnston, M., Hillier, L., Riles, L., Albermann, K., Andre, B.,  
Ansoyge, W., Benes, V., Bruckner, M., Dellus, H., Dubois, E.,  
Dumesthorff, A., Entian, K.D., Floeth, M., Goffeau, A., Hebling, U.,  
Heumann, K., Heuss-Neitzel, D., Hilbert, H., Hilger, F., Kleine, K.,  
Kotter, P., Louis, E.J., Messenguy, F., Mewes, H.W., Miosga, T.,  
Mostl, D., Muller-Auer, S., Nentwich, U., Obermaier, B., Piravandi, E.,  
Pohl, T.M., Portetel, D., Purnelle, B., Rechmann, S., Rieger, M.,  
Rinke, M., Rose, M., Scharfe, M., Scherens, B., Scholler, P.,  
Schwager, C., Schwarz, S., Underwood, A.P., Urrestazu, L.A.,  
Vandenbol, M., Verhasselt, P., Vierendeels, F., Voet, M., Voickaert, G.,  
Voss, H., Wambutt, R., Wedler, E., Wedler, H., Zimmermann, F.K.,  
Zollner, A., Hani, J. and Hoheisel, J.D.

The nucleotide sequence of Saccharomyces cerevisiae chromosome XII  
Nature 387 (6632 Suppl), 87-90 (1997)

97313267

9169871

2 (bases 1 to 38868)

Pauley, A.

The sequence of S. cerevisiae cosmid 8167

Unpublished (1994)

3 (bases 1 to 38868)

Waterston, R.

Direct Submission

Submitted (16-SEP-1994)

4 (bases 1 to 38868)

Cherry, J.M. ISSION

Direct Submission

Submitted (22-AUG-1997)

Sequenced by: Saccharomyces Genome Database

Genome Sequencing Center

Department of Genetics, Washington University,

St. Louis, MO 63110, USA

Curated by: Saccharomyces Genome Database

URL: <http://genome-www.stanford.edu/>

e-mail: [yeast-curator@genome.stanford.edu](mailto:yeast-curator@genome.stanford.edu)

Location/Qualifiers

1..38868

/organism="Saccharomyces cerevisiae"

/db\_xref="taxon:4932"

/map="XII"

/strain="S288C (AB972)"

<2..58

<2..58

<2..58

<2..58

<2..58

<2..58

<2..58

<2..58

<2..58

<2..58

<2..58

<2..58

<2..58

<2..58

<2..58

<2..58

<2..58

FEATURES

source

gene

CDS

gene

CDS

note="Weakly similar to protein of unknown function from

Halobacterium halobium (Swissprot accession number

P13043)"



Db 30307 ATGTTCTAGTAATAATCTAAAAAGTTGCTTACTGTAAAAACATCAATTTGGCTGGGTTCAGTGAC 30366  
Qy 196 tgcattgccttatgagaactcccaaatcaacaagactctggactccgcttttgatgaact 255  
Db 30367 AGCATGTGCCTATGAGAAATCCAAATCTACAAAACACTAGACAGCGCCTTAATGAAGTT 30426  
Qy 256 tgcacgccaatgctcaagttatcaaggtttacacactggagacatgaagaacatctacct 315  
Db 30427 AGCATCCCAATGTTCAAGCATCAAGTTTATACCTTTAGAGGACATGAAGAATATTTATTT 30486  
Qy 316 taatgaagttaactactcctgcgctcctgagaaatccgataaagaagacagttgtttcaca 375  
Db 30487 AAATGCGTCAAAATTTATTTAGAGCACCCTGAGAAAAGTATTAATAAAACCGGTGTATGCA 30546  
Qy 376 accgttgatggcaaatgagacgacctatcactactactactatgagaaactatggatcca 435  
30547 ACCGCTCATGGCAACGAGACAGGCTATCATTTATATATGAGGAAAATATATGATATCCA 30606  
436 cttgaatttgatcgatctcaatggtgcgcatggggcctcgtctcttcttctgggtcgagtt 495  
30607 TCTTAACCTAATGGCTCTCAATGGTGGCTTGGGTCTCGTCTTCTTCTGGGTGGCTGT 30666  
Qy 496 ccttacgcgcgaactatttgaacattctcaaacgcgctattcggcaagaacattatggc 555  
Db 30667 GCTTACTGCAAGCCACTATCTGAAACATTTCTGAAAAGGGTGTGGTGAAGAACATCATATGGC 30726  
Qy 556 aaattctgttaagaagtccttactaccacaagcgctttacaagaactacacagagagaac 615  
Db 30727 AACTCCGTCNAANAATCACTTATTTATCCTCTGTTTACANAAGATTTATATGAACGAAC 30786  
Qy 616 ttcttatctttggaaacgctttggccactcaactttacaaactcagagagaaagactcgtagt 675  
Db 30787 TTTTATTTATGGAAGGCTCTACCATTTAAATTTTACAACTCGAGGCAAGGCTCTCGTGT 30846  
Qy 676 tcttatcttgctcatcttgactattctcactctctcactctcttctcggacataacatcaagtggc 735  
Db 30847 ATTAATTTTGTATTTGACTATATTTATCTCTCAGTTTGGTCAATATATTAACATTC 30906  
Qy 736 acatctctacgatagactagatgagaagatacaatggcattcgtctcagccgctgctga 795  
Db 30907 ACACCCATATGATAGGCCAGATGGAAGNAGTATGGCCTTTGTGAGTCGTAGAGAGA 30966  
Qy 796 cttgatggcaatgcctctttcccccgttggtgtaccttttogggtatccggaaacacccctt 855  
Db 30967 CTTGATGCCATTCGACTTTTCCAGTAGTCTATCTATTTCGGAATAAGAAATAATCCCTT 31026  
856 catcccaatcacggatgagctttagcttacttcaacttttaccacaaatggtcagcata 915  
31027 CATCCCTATACAGGGCTTTCCCTTTTCTACATTTAAATTTCTATCAATAAATGGTCTGGCTA 31086  
Qy 916 cgtctgcttcattgtagccgtcgctccattcaatcgcttatgacccgttcaggagttaacg 975  
Db 31087 CGTTTGTTCATGTTGGCCGTGTACACTCAATTTGTATGACCCGCTCGGAGTGAAAAG 31146  
Qy 976 aggsagttaccagtcctcttgaagaaattctactctcagatggggaatagtagccacaat 1035  
Db 31147 AGGTGTGTTTCAAAAGTCTGGTTAGGAAATTTACTTTAGGTGGGTATAGTGGCAACGAT 31206  
Qy 1036 tcttatgtccatcatctttccagtcggaagaggtcttcgaagaccgaggttatgaaat 1095  
Db 31207 ATTAATGCTATTTATTTTCCAAAAGTAAAGATTTTGAAGAAATAGAGGGTATGAGAT 31266  
Qy 1096 cttcttacttattcacaaagccatgaacatcatgttttatcatagctatgtattaccattg 1155  
Db 31267 ATTCTCTTATTCTAAGCGGATGAATATATGTTTCAATTATGCCATGTACTACCATTTG 31326  
Qy 1156 ccacacactaggatggatgggctggatctgtgtccatggctggcctcctctgcttcgacag 1215  
Db 31327 TCACACCCCTGGCTGGATGGGTGGATTTGGTCAATGGCTGGTATTTATGCTTTGTATAG 31386  
Qy 1216 gttctgcgaattgtacgatcatcatgaacggaggtcttaagaccgccaactttgtcgac 1275  
Db 31387 ATTCTCGAGGATTCTTGAATAATCATGAATGGTGGCTTGAATAACTGCTACTTTGAGTAC 31446

Qy 1276 cacagatgatttaacgcttatcaagatctctgtcaagaagcctcaagttcttctcaagttatca 1335  
Db 31447 CACTGATGATTCTAATGTTATTAATAATTTCAAGTAAAAAACCAAAAGTTTTTCAAGTACCA 31506  
Qy 1336 agtgggagcatttgcctatatgtacttcttcttccacaaaaacagcctgggttctctacagttt 1395  
Db 31507 AGTAGGAGCTTTTCATACATACATGTTATTTCTTATCACCAAAAAAGTCATGGTTCTTATAGTTT 31566  
Qy 1396 tcaatctcatcctctcacagtcctcatcagaagaagcagagatcctcaacaaccagatca 1455  
Db 31567 CCAATACATCCATTTACAGTATTATCGGAACGACACCGTGATCCAAACATTCAGATCA 31626  
Qy 1456 actaactatgtacgtcaaaagcctaagaaggcattacagagagctactctttagcaagttctt 1515  
Db 31627 ATTGACGATGTACGTAAAGGCAATAAAGGTATCACTCGAGCTTTTGTATCGAAAGTTCT 31686  
Qy 1516 aagcgtctcaaacatcacctgttgattgcaagaatttcttagaggagaccatagtcggttaac 1575  
Db 31687 AAGTGCTCCAAATCATATGTTGATTGTGTAATAATATTCCTTGAAGGCCCATATGGTGTAAAC 31746  
Qy 1576 tgtccctcacattgccaaaacttaagagaaatcttagtaggagtagctcgccctcgccgt 1635  
Db 31747 GGTTCACATATCCTAAGCTAAAAAGAAATCTGTGAGGTAGCCGTGGTGGGTGT 31806  
Qy 1636 ggcagcattctacccccatttcctagaaatgcttagattgcttagcactgacatgacatgca 1695  
Db 31807 TGGCGCTATTTATCCGCATTTTGTGCAATGTTTACGGTTTACCATCTACTGATCAACTTCA 31866  
Qy 1696 gcaaatgtctactggatcgctcaacgaccttagtcaccccttaagtggttcgaaaaagcagct 1755  
Db 31867 GCATAAAATTTTACTGGATTTGTTAATGACCTATCCCATTTGAAATGGTTGAAATGAAT 31926  
Qy 1756 acaatgcttaagagaaaatctgtgaagctctgtcatctcacactgggtcatcagttga 1815  
Db 31927 GCAATGGTTAAAGAGAAAAGTTGTGAAGTCTCAGTCAATATATCTGTTCCAGTGTGTA 31986  
Qy 1816 ggaacaaactcagatgagtcacttaagggtttctcgatgacaagaagaatctgaaatcac 1875  
Db 31987 GGACACAAATTCAGATGAGAGTACAAAAGTTTGTGATGATAAAGAAACGAAATCAC 32046  
Qy 1876 cgtgaatgccttaacaagagccagacctcaagagctcaagagctagtagagatcagagatacaatt 1935  
Db 32047 TGTGTAATGTCTCAATAAAAGACCTGATTTGAAAAGAACCTAGTGGCTCGGAAATAAAACT 32106  
Qy 1936 gtcagaactcagacaacaacatcacttctactcatcgagaccagcagacttcaatga 1995  
Db 32107 CTCAGAACTAGAGAAATAATATATTTACCTTTTATTCCTGCGGGCAGCAACGTTTAACGA 32166  
Qy 1996 cgaacttaggaatgcagttgtacaaggtatcgattcttagtctgaagatagatgtcgaaact 2055  
Db 32167 CGATTTTAGAAATGCAGTGGTCCAAGGTATAGACTCTTCTCTTGAAGATGACGTTGAACT 32226  
Qy 2056 agagagagaggttttactctggttaagagctcaagct 2091  
Db 32227 AGAAGAAAGATTTTACATGGTAAGGCCCTCTGTT 32262

## RESULT 4

LOCUS A86082 694 bp DNA PAT 21-JAN-2000  
DEFINITION Sequence 741 from Patent EP0866129.  
ACCESSION A86082  
VERSION A86082.1 GI:6734681  
KEYWORDS  
SOURCE Eremothecium gossypii.  
ORGANISM Eremothecium gossypii  
Eukaryote; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
Saccharomycetales; Eremotheciaceae; Eremothecium.  
REFERENCE  
1 (bases 1 to 694)  
AUTHORS Mohr C. and Knechtie, P.  
TITLE Genomic DNA sequences of *Ashbya gossypii* and uses thereof  
JOURNAL Patent: EP 0866129-A 741 23-SEP-1998;

FEATURES		CIBA GEIGY AG (CH)		Location/Qualifiers	
source		1..694		/organism="Eremothecium gossypii"	
BASE COUNT		164 a 154 c 189 g 187 t			
ORIGIN					
Query Match 7.1%; Score 147.6; DB 6; Length 694;					
Best Local Similarity 54.6%; Pred. No. 4.7e-32;					
Matches 344; Conservative 0; Mismatches 274; Indels 12; Gaps 2;					
QY	1006	ctacttcagatgggaataagtaggcacacaattcttatgtccatcatcattttccagtcgaa	1065		
DB	76	CTATTTTAAATGGGGTGTGTTGCAACGGTTTGGCCGGGCTCTTGCTCGCATAGTGA	135		
QY	1066	gaaggtcttcagaaacccagggttatgaaatcttcttacttattcacaagccatgaacat	1125		
DB	136	AAAGTTTCATTCGCCAAGAAGACGTACGAATCTTCTGTATATGCACAAGCTCTTCAACAT	195		
QY	1126	catgtttatcatagctatgtattaccattgcccacacataggatggatggctgactcg	1185		
DB	196	TGCTTTTATGTATGATGTATATGACATCAAAACGCTGGGATGGCAGCGCTGGGCTCG	255		
QY	1186	gtccatggctggcatcctctgtcttcacaggttctgcgaattgtacgtatcatcatgaa	1245		
DB	256	GTGCATGGTGGCCATCTACTGCTTCGAGCGTGTGGCCGGATAGCTCGCATGTACTTGC	315		
QY	1246	cggaggtcttaagaccgcacattgtcgaccacagatgattctaaagttatcaagatctc	1305		
DB	316	TGGAGGCATCAAGAAGCCACATTT---AACAGATGTTGGGATCGCGTCTCAAGATGAC	372		
QY	1306	tgtcaagaagcctaagttcttcaagtatcaagtggagacattggcctatatgtactttct	1365		
DB	373	AGTGGAGAAGCCAAAGCATTTCAAATATTACCGGGGGCTTATGTTTTCGTTTATTTAT	432		
QY	1366	ttcaccaaaatcagcctgtgttctacagtttcaatctcctcccttcacagtcctatcaga	1425		
DB	433	TAGTGGGAAGGATGCTTGGTTCTATCCATTCAGTCGACCCGTTCCACCGTCTTAATAC	492		
QY	1426	aaggcacagagatccttaaacaccacagatcaactaaatagtacgtcaaaagtaacaagg	1485		
DB	493	ACCCAAAGATCGATGGCGACAACCTGGTG-----ATTATTCAAAGTGCCACAAGGG	543		
QY	1486	cattacagagtagtactcttagcaaaagtcttaagcgtcccaaccataccgttgattgcaa	1545		
DB	544	CGTGACGACGACGCTGCTAAACAGGATCTTCTATCCGGGAAAGAGTCCATCGAATACAA	603		
QY	1546	gattttcttagagggaccatatggcgtaactgtccctcacattgcccacacttaagagaaa	1605		
DB	604	GGTGCTTCTAGAAGGCCCTATGGAACACCATTCGCGGGCTTGCTGCTCCTGACCGGG	663		
QY	1606	tctagtaggagtagtgcgggcctcgcgct	1635		
DB	664	CTACGTGGGCGCCAGCGCAGGTCTTGGCGT	693		
RESULT 5					
AR155575					
LOCUS					
DEFINITION					
Sequence 741 from patent US 6239264.					
ACCESSION					
AR155575					
VERSION					
AR155575.1 GI:15123628					
KEYWORDS					
SOURCE					
Unknown.					
ORGANISM					
Unknown.					
REFERENCE					
1 (bases 1 to 694)					
AUTHORS					
Philippsen, P., Pohlmann, R., Steiner-Lange, S., Mohr, C., Wendland, J., Knechtie, P., and Reibischung, C.					
TITLE					
Genomic DNA sequences of ashbya gossypii and uses thereof					
JOURNAL					
Patent: US 6239264-A 741 29-MAY-2001;					

FEATURES		Location/Qualifiers	
Source		1..694	
BASE COUNT		164 a 154 c 189 g 187 t	
ORIGIN			
Query Match 7.1%; Score 147.6; DB 6; Length 694;			
Best Local Similarity 54.6%; Pred. No. 4.7e-32;			
Matches 344; Conservative 0; Mismatches 274; Indels 12; Gaps 2;			
QY	1006	ctacttcagatgggaataagtaggcacacaattctttagtcacatcatcattttccagtcoga	1065
DB	76	CTATTTTAAATGGGGTGTGTTGCAACGGTTTGGCCGGGCTCTTGCTCGCATAGTGA	135
QY	1066	gaaggtcttcagaaacccagggttatgaaatcttcttacttattcacaagccatgaacat	1125
DB	136	AAAGTTTCATTCGCCAAGAAGACGTACGAATTTCTTCTGTATATGCACAAGCTCTTCAACAT	195
QY	1126	catgtttatcatagctatgtattaccattgcccacacataggatggatggctgactctg	1185
DB	196	TGCTTTTATGTATGATGTATATGACATCAAAACGCTGGGATGGCAGCGCTGGGCTCG	255
QY	1186	gtccatggctggcatcctctgtcttcacaggttctgcgaattgtacgtatcatcatgaa	1245
DB	256	GTGCATGGTGGCCATCTACTGCTTCGAGCGGTGTGGCCGGATAGCTCGCATGTACTTGC	315
QY	1246	cggaggtcttaagaccgcacattgttcgaccacagatgattctaaagttatcaagatctc	1305
DB	316	TGGAGGCATCAAGAAGCCACATTT---AACAGATGTTGGGATCGCGTCTCAAGATGAC	372
QY	1306	tgtcaagaagcctaagttcttcaagtatcaagtgggagacattggcctatatgtactttct	1365
DB	373	AGTGGAGAAGCCAAAGCATTTCAAATATTACCGGGGGCTTATGTTTTCGTTTATTTAT	432
QY	1366	ttcaccaaaatcagcctgttctcagtttcaatctcctcccttcacagtcctatcaga	1425
DB	433	TAGTGGGAAGGATGCTTGTTCTATCCATTCAGTCGACCCGTTCCCGTCTCTTAATAC	492
QY	1426	aaggcacagagatccttaaacaccacagatcaactaaatagtacgtcaaaagtaacaagg	1485
DB	493	ACCCAAAGATCGATGGCGACAACCTGGTG-----ATTATTTCAAAGTGCCACAAGGG	543
QY	1486	cattacgagagtagtctttagcaaaagtcttaagcgtcccaaccataccgttgattgcaa	1545
DB	544	CGTGACGACGACGCTGCTAAACAGGATCTTCTATCCGGGAAAGAGTCCATCGAATACAA	603
QY	1546	gattttcttagagggaccatatggcgtaactgtccctcacattgcccacacttaagagaaa	1605
DB	604	GGTGCTTCTAGAAGGCCCTATGGAACACCATTCGCGGGCTTGCTGCTCCTGACCGGG	663
QY	1606	tctagtaggagtagtgcgggcctcgcgct	1635
DB	664	CTACGTGGGCGCCAGCGAGGTCTTGGCGT	693
RESULT 6			
LOCUS	E66100	594 bp	DNA
DEFINITION	Genome DNA sequence of Ashbya gossypii and use thereof.		
ACCESSION	E66100	PAT 07-FEB-2001	
VERSION	E66100.1	GI:13020441	
KEYWORDS	JP 1999225770-A/740.		
SOURCE	unidentified.		
ORGANISM	unidentified.		
REFERENCE	1 (bases 1 to 694)		
AUTHORS	Peter, P.R.P., Steiner, C.M., Juergen, W.P.K.K. and Reibishun.		
TITLE	Genome DNA sequence of Ashbya gossypii and use thereof		
JOURNAL	Patent: JP 1999225770-A 740 24-AUG-1999;		
COMMENT	NOVARTIS AG		
	OS	PAG1482RP	
	PN	JP 1999225770-A/740	





```
Db 2508 GTGCAAGAAAAGCAACCGTCTCTATTAAAGGTGACGATAC-----TTGAAGATTGAGG 2561
Qy 1308 tcaagaagcctaagttcttcaagtatcaagtggagagcttggctatatgtactttttt 1367
Db 2562 TTCCCAACACCTAAATATTGGAAGTCGGTAGCAGGTGGACATGCTTTTATTTCATTCTTGA 2621
Qy 1368 caccataatca 1378
      |||||
Db 2622 AACCAACCTTA 2632

RESULT 9
LOCUS A85974 711 bp DNA PAT 21-JAN-2000
DEFINITION Sequence 633 from Patent EP0866129.
ACCESSION A85974
VERSION A85974.1 GI:6734573
KEYWORDS
ORGANISM Eremothecium gossypii.
REFERENCE 1 (bases 1 to 711)
AUTHORS Mohr,C. and Knechtle,P.
TITLE Genomic DNA sequences of Ashbya gossypii and uses thereof
JOURNAL Patent: EP 0866129-A 633 23-SEP-1998;
CIBA GEIGY AG (CH)
FEATURES
    source
        1. 711
            /organism="Eremothecium gossypii"
            /db_xref="taxon:33169"
BASE COUNT 165 a 193 c 167 g 186 t
ORIGIN

Query Match 4.3%; Score 90.2; DB 6; Length 711;
Best Local Similarity 50.3%; Pred. No. 3.8e-15;
Matches 280; Conservative 0; Mismatches 268; Indels 9; Gaps 2;

Qy 1526 aaccataccgttgattgcaagattttcttagaggagaccataatggcgttaactgtccctcac 1595
      |||
Db 17 AAAGAGTCCATCGAATACAAAGTGCTTCTTAGAAGGGCCCTATGGAACACACCATTTCCGGG 76
Qy 1586 attgccaaaactaagagaactctagtaggagtagctgcgggcctcgcggtggcagccatc 1645
      |||
Db 77 CTTCGCTCTCTGACGGCGCTACGTGGGCCGCCAGCGAGGTCTTTGGCGTAGCAGCGGTC 136
Qy 1646 taccctcattctgtagaatgccttagattgcctagcactgatcaactgcagcacaagtctc 1705
      |||||
Db 137 TACCACACATTCTGT-----CTCTCTGTGGACAAGGAAGCCAGTTACCCCATTCATTTC 190
Qy 1706 tactggatcgtaacgaccttagtcaccttaagtgggttcgaaacgagactacaatggctt 1765
      |||||
Db 191 TACTGGATTATAAATGACCTTTTCATATCTGCATTGGTTTTCGCATGAGCTGAGGTACCTG 250
Qy 1766 aaggagaaactctgtgaagtctctgcactcacactgggtcatcagtgaggagatacaaac 1825
      |||||
Db 251 GCGGACCGGAACCTCGACATCAAAATTTATTACAGGAGGACCAATGATGTCGGCTAAAGAA 310
Qy 1826 tcagatgagtcacctaagggttttcg---atgacaagaagaagtctgaaatcacccgtagaa 1882
      |||||
Db 311 CTGACCCACAGATGTGCGGATTCGCCCTCTCGGAAGTTTCGGGATTTCGCTGGATATCTGC 370
Qy 1883 tgccttaacaagaggcgaacactcaagagctagtggatcagagatacaaatgtccgaa 1942
      |||||
Db 371 AGGCTCTCTCTGCGCCCGAGATCTCAAAAGAGATCGTGGAAAGAGACGATCCTCTCTCGTCT 430
Qy 1943 ctcgagaacaacaacatcacttctactcatgcgagaccgagcactttcaatgacgacttt 2002
      |||||
Db 431 AACACGGCACAGGACGTACCGTTTATTAGCAGCGGCCCTTCGACCTTTTAAATGACCATTTTC 490
Qy 2003 aggaatgcagttgtacaaggtatcgatttcttagtctgaagatagatgtcgaactagaggag 2062
      |||||
Db 491 CGCTATGCTGTGAAATCTAGCATCACGGGCAAACTCCACGTGTGTGATGTCGACCTTAGAGGAG 550
Qy 2063 gagagtttacttgta 2079
      |||||
Db 551 GAAAGCTACACCTGGTA 567
```

```
Db 491 CGCTATGCTGTGAAATCTAGCATCACGGGCAAACTCCAGTGTGTGATGTCGACCTTAGAGGAG 550
Qy 2063 gagagtttacttgta 2079
      |||||
Db 551 GAAAGCTACACCTGGTA 567

RESULT 10
LOCUS ARI55467 711 bp DNA PAT 08-AUG-2001
DEFINITION Sequence 633 from patent US 6239264.
ACCESSION ARI55467
VERSION ARI55467.1 GI:15123520
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 711)
AUTHORS Philippsen,P., Pohlmann,R., Steiner-Lange,S., Mohr,C., Wendland,J.,
Knechtle,P. and Reibschung,C.
TITLE Genomic DNA sequences of ashbya gossypii and uses thereof
JOURNAL Patent: US 6239264-A 633 29-MAY-2001;
FEATURES
    source
        1. 711
            /organism="unknown"
BASE COUNT 165 a 193 c 167 g 186 t
ORIGIN

Query Match 4.3%; Score 90.2; DB 6; Length 711;
Best Local Similarity 50.3%; Pred. No. 3.8e-15;
Matches 280; Conservative 0; Mismatches 268; Indels 9; Gaps 2;

Qy 1526 aaccataccgttgattgcaagattttcttagaggagaccataatggcgttaactgtccctcac 1595
      |||
Db 17 AAAGAGTCCATCGAATACAAAGTGCTTCTTAGAAGGGCCCTATGGAACACACCATTTCCGGG 76
Qy 1586 attgccaaaactaagagaactctagtaggagtagctgcgggcctcgcggtggcagccatc 1645
      |||
Db 77 CTTCGCTCTCTGACGGCGCTACGTGGGCCGCCAGCGAGGTCTTTGGCGTAGCAGCGGTC 136
Qy 1646 taccctcattctgtagaatgccttagattgcctagcactgatcaactgcagcacaagtctc 1705
      |||||
Db 137 TACCACACATTCTGT-----CTCTCTGTGGACAAGGAAGCCAGTTACCCCATTCATTTC 190
Qy 1706 tactggatcgtaacgaccttagtcaccttaagtgggttcgaaacgagactacaatggctt 1765
      |||||
Db 191 TACTGGATTATAAATGACCTTTTCATATCTGCATTGGTTTTCGCATGAGCTGAGGTACCTG 250
Qy 1766 aaggagaaactctgtgaagtctctgcactcacactgggtcatcagtgaggagatacaaac 1825
      |||||
Db 251 GCGGACCGGAACCTCGACATCAAAATTTATTACAGGAGGACCAATGATGTCGGCTAAAGAA 310
Qy 1826 tcagatgagtcacctaagggttttcg---atgacaagaagaagtctgaaatcacccgtagaa 1882
      |||||
Db 311 CTGACCCACAGATGTGCGGATTCGCCCTCTCGGAAGTTTCGGGATTTCGCTGGATATCTGC 370
Qy 1883 tgccttaacaagaggcgaacactcaagagctagtggatcagagatacaaatgtccgaa 1942
      |||||
Db 371 AGGCTCTCTCTGCGCCCGAGATCTCAAAAGAGATCGTGGAAAGAGACGATCCTCTCTCGTCT 430
Qy 1943 ctcgagaacaacaacatcacttctactcatgcgagaccgagcactttcaatgacgacttt 2002
      |||||
Db 431 AACACGGCACAGGACGTACCGTTTATTAGCAGCGGCCCTTCGACCTTTTAAATGACCATTTTC 490
Qy 2003 aggaatgcagttgtacaaggtatcgatttcttagtctgaagatagatgtcgaactagaggag 2062
      |||||
Db 491 CGCTATGCTGTGAAATCTAGCATCACGGGCAAACTCCACGTGTGTGATGTCGACCTTAGAGGAG 550
Qy 2063 gagagtttacttgta 2079
      |||||
Db 551 GAAAGCTACACCTGGTA 567
```







Oy 272 agtatcaaggtttacacactgga 294  
|||||  
Db 23 AGTATCAAGGTTTACACACTGGA 1

Search completed: May 14, 2002, 07:24:06  
Job time: 11053 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run On: May 14, 2002, 04:19:48 ; Search time 1786.89 Seconds  
(without alignments)  
12580.611 Million cell updates/sec

Title: US-09-646-825-1

Perfect score: 2092

Sequence: 1 gaattcttagactccacca.....cttggtaagactcaagctt 2092

Scoring table:

IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 11351937 seqs, 5372889281 residues

Total number of hits satisfying chosen parameters: 22703874

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*  
1: em\_estfun:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estom:\*  
5: em\_estpl:\*  
6: em\_estba:\*  
7: em\_estro:\*  
8: em\_estov:\*  
9: em\_htc:\*  
10: gb\_estl:\*  
11: gb\_est2:\*  
12: gb\_htc:\*  
13: gb\_gss:\*  
14: em\_gss\_fun:\*  
15: em\_gss\_hum:\*  
16: em\_gss\_inv:\*  
17: em\_gss\_pln:\*  
18: em\_gss\_pro:\*  
19: em\_gss\_rod:\*  
20: em\_gss\_vrt:\*  
21: em\_gss\_other:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
c 1	187.4	9.0	962	13	CNS060YA
c 2	149	7.1	515	13	A2931275
c 3	71.4	3.4	966	13	AL422333
c 4	69	3.3	968	13	CNS06HK9
c 5	53.4	2.6	1015	13	CNS06VLD
c 6	50.6	2.4	958	13	CNS06WLF
c 7	48.2	2.3	851	13	CNS079Q2
c 8	47.4	2.3	987	13	CNS06KPL
c 9	44.2	2.1	1101	13	CNS01705
c 10	44	2.1	805	13	AQ858562
c 11	42.8	2.0	888	13	A2528430
c 12	42.8	2.0	949	13	CNS06NMI

13	42.8	2.0	1101	13	CNS0039G
14	41	2.0	1043	13	CNS07DHO
15	41	2.0	2149	11	EG714799
16	40.8	2.0	391	10	AV677876
17	40.6	1.9	843	13	AZ551618
18	40.6	1.9	905	13	AZ550256
19	40.2	1.9	908	13	AZ548467
c 20	40	1.9	891	13	AZ683582
c 21	39.6	1.9	300	10	AI893200
c 22	39.6	1.9	416	10	AI483599
c 23	39.6	1.9	487	13	AZ111934
c 24	38.6	1.9	880	13	AZ669474
c 25	38.8	1.9	320	10	BB139106
c 26	38.8	1.9	974	13	CNS06O5J
c 27	38.8	1.9	1101	13	CNS00KDX
c 28	38.6	1.8	633	10	AU214668
c 29	38.6	1.8	816	13	AZ535744
c 30	38.2	1.8	1630	12	AK009327
c 31	37.8	1.8	328	13	AZ923069
c 32	37.6	1.8	287	10	BB370647
c 33	37.6	1.8	290	11	EG937781
c 34	37.6	1.8	1272	11	EG394360
c 35	37.4	1.8	369	13	AQ152254
c 36	37.4	1.8	505	10	AI333640
c 37	37.4	1.8	549	10	AL500267
c 38	37.2	1.8	584	11	EG463815
c 39	37.2	1.8	621	10	AU203061
c 40	37.2	1.8	716	10	AU213521
c 41	37	1.8	519	13	AZ803420
c 42	37	1.8	668	13	AQ454443
c 43	37	1.8	944	12	AK015703
c 44	36.6	1.7	849	13	AZ546009
c 45	36.4	1.7	321	11	EG224888

#### ALIGNMENTS

RESULT 1  
CNS060YA/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

REFERENCE

AUTHORS

CNS060YA 962 bp DNA GSS 05-JUL-2001  
T7 end of clone AV0AA010B01 of library AV0AA from strain CBS 379 of  
Saccharomyces exiguus, genomic survey sequence.

AL408680

AL408680.1 GI:12175485

GSS.

Saccharomyces exiguus.

Saccharomyces exiguus

Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;

Saccharomycetales; Saccharomycetaceae; Saccharomycetes.

1 (bases 1 to 962)

Bon,E., Neuveglise,C., Lepingle,A., Wincker,P., Artiguenave,F.,

Gaillardin,C. and Casaregola,S.

Genomic exploration of the hemiascomycetous yeasts: 6.

Saccharomyces exiguus

FEBS Lett. 487 (1), 42-46 (2000)

20584716

2 (bases 1 to 962)

Souci,J., Aigle,M., Artiguenave,F., Blandin,G.,

Bolotin-Fukuhara,M., Bon,E., Brottier,P., Casaregola,S.,

de-Montigny,J., Dujon,B., Durrens,P., Lepingle,A., Llorente,B.,

Malpertuy,A., Neuveglise,C., Ozier-Kalogeropoulos,O., Potier,S.,

Saurin,W., Tekala,F., Toffano-Nioche,C., Wesolowski-Louvel,M.,

Wincker,P. and Weissenbach,J.

Genomic exploration of the hemiascomycetous yeasts: 1. A set of

yeast species for molecular evolution studies

FEBS Lett. 487 (1), 3-12 (2000)

20584711

3 (bases 1 to 962)

Genoscope.

Direct Submission

Submitted (07-SEP-2000) Genoscope - Centre National de Sequencage,

2 rue Gaston Cremieux, CP 5706, 91057 EVRY cedex, FRANCE. (E-mail :



Qy 840 tccggaacaccccttcacccaatcaccgaggttagctttagtactttcaactttacc 899  
Db 129 TCAGAAACAATCTTTTCATGTATTAACTGGCTTTCCCTTTACCACATCTACTATTACC 188  
Qy 900 acaaatggtcagcacaacgtctgcttcattgtagccgctgctccattcaatcttatgaccg 959  
Db 189 ATAGATGCTGCTTACATTTGTACTGTATTGGCATTCATTCAATCAATATTATGGACAA 248  
Qy 960 ctccaggagttaaacgagaggtattccagtcctcttgtaagaataattctactcagatggg 1019  
Db 249 GCTATGCAACGAATGGCCATTTTATACAGAAATGACTGAGATTAATTTATGGAGATGGG 308  
Qy 1020 gaatagtagcaccaattctatgtccatcatctatttccagtcgagaggtcttcagga 1079  
Db 309 GGATTGTTGCCACCACTTTGTTATTCTTACTGATATTTCAGCTGAAAGACAATCAGAA 368  
Qy 1080 accgaggttatgaattcttactattatccaaagccatgaacatcatgtttatcatag 1139  
Db 369 ATCTAATGTATGAGTATTTTAAATATATACATCACTTTTATAGTATTATATTATTGTTG 428  
1140 ctatgtattaccattgccacacactagatgagtggtgctgaatctggtccatggtgcca 1199  
Db 429 CAATGTATCACCATGTACGAAGATTGTTGGATTGGTTGGATTGGTCGATGGCTGTA 488  
Qy 1200 tectctgcttoga 1212  
Db 489 TATATGGATTGA 501  
RESULT 3  
LOCUS CNS062HJ 966 bp DNA GSS 06-JUL-2001  
DEFINITION T3 end of clone XAY0AA002F02 of library XAY0AA from strain CBS 6340  
of Kluyveromyces thermotolerans, genomic survey sequence.  
ACCESSION AL422333  
VERSION AL422333.1 GI:12205527  
KEYWORDS GSS.  
SOURCE Kluyveromyces thermotolerans.  
ORGANISM Kluyveromyces thermotolerans.  
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
Saccharomycetales; Saccharomycetaceae; Kluyveromyces.  
REFERENCE 1 (bases 1 to 966)  
AUTHORS Malpertuy,A., Liorente,B., Blandin,G., Artiguenave,F., Wincker,P.  
and Dujon,B.  
TITLE Genomic exploration of the hemiascomycetous yeasts: 10.  
JOURNAL Kluyveromyces thermotolerans  
MEDLINE FEBS Lett. 487 (1), 61-65 (2000)  
REFERENCE 20584720  
AUTHORS Souciet,J.L., Aigle,M., Artiguenave,F., Blandin,G.,  
Bolotin-Fukuhara,M., Bon,E., Brottier,P., Casaregola,S.,  
de-Montigny,J., Dujon,B., Durrens,P., Lepingle,A., Liorente,B.,  
Malpertuy,A., Neuveglise,C., Ozier-Kalogeropoulos,O., Potier,S.,  
Saurin,W., Tekala,F., Toffano-Nioche,C., Wesolowski-Louvel,M.,  
Wincker,P. and Weissenbach,J.  
TITLE Genomic exploration of the hemiascomycetous yeasts: 1. A set of  
yeast species for molecular evolution studies  
JOURNAL FEBS Lett. 487 (1), 3-12 (2000)  
MEDLINE 20584711  
REFERENCE 3 (bases 1 to 966)  
AUTHORS Genoscope.  
TITLE Direct Submission  
JOURNAL Submitted (08-SEP-2000) Genoscope - Centre National de Sequencage,  
2 rue Gaston Cremieux, CP 5706, 91057 EVRY cedex, FRANCE. (E-mail :  
seqrefgenoscope.cns.fr - Web : www.genoscope.cns.fr)  
COMMENT This GSS is part of a random genomic sequencing program of thirteen  
yeast species: Saccharomyces bayanus var. uvarum, Saccharomyces  
exiguus, Saccharomyces servazzii, Zygosaccharomyces rouxii,  
Saccharomyces kluyveri, Kluyveromyces thermotolerans, Kluyveromyces  
lactis var. lactis, Kluyveromyces marxianus var. marxianus, Pichia  
angusta, Debaryomyces hansenii var. hansenii, Pichia sorbitophila,  
Candida tropicalis and varrowia lipolytica. Genomic inserts of 3 to  
5 kb were prepared and both extremities were sequenced. See

keywords for description of this sequence and for the sequence of  
the other extremity of this insert.  
FEATURES  
source 1..966  
/organism="Kluyveromyces thermotolerans"  
/strain="CBS 6340"  
/db\_xref="taxon:4916"  
/clone="XAY0AA002F02"  
/clone\_lib="XAY0AA"  
/note="end : T3"  
/note="similar to Saccharomyces cerevisiae ORF YLR215c [  
strong similarity to rat cell cycle progression related  
D123 protein ]"  
/evidence=not\_experimental  
misc\_feature complement(<388..>938)  
/note="similar to Saccharomyces cerevisiae ORF YLR214w [  
FREL ; ferric (and cupric) reductase ]  
1 putative frameshift(s)"  
/evidence=not\_experimental  
BASE COUNT 273 a 218 c 212 g 262 t 1 others  
ORIGIN  
Query Match 3.4%; Score 71.4; DB 13; Length 966;  
Best Local Similarity 49.9%; Pred. No. 8.9e-09;  
Matches 317; Conservative 0; Mismatches 306; Indels 12; Gaps 5;  
Qy 1446 accagatcaactaactatgtactcaaaagctaaagggcattacgagagtactctta 1505  
Db 937 ACGGAACAGAGATGCGAGTGTACTTCAAAGCTCAGAAAGGTGTACCGGGCACATGCTCT 878  
Qy 1506 gcaaatgtcttaagcgtctcaaacaccataccgttgattgaagattttcttagagggaacat 1565  
Db 877 ATAAAGCTTCTCAGAAAGTGGCGAGAAT--CCATAAAACATAAGAGATCATGGTTGAGGGCCTT 820  
Qy 1566 atggcgtaactgtccctcacattgccaaacttaagagaaatctagtagtagtgcgcg 1625  
Db 819 ACGGTGACACCTTGCTCAGATCGCCAAAGTGAATAGGAGAATTTGTTGTCGCTGCTG 760  
Qy 1626 gctcggcggtggcagccatctaccctccatttcgttagaatgccttagatgccttagcactg 1685  
Db 759 GGTGGGTGTGACAGCGCTTTACGCTCAAAATTTCTCCACCTGTAGAAAGCACAGGTACCCG 700  
Qy 1686 atcaactgcagcacaagtcttactgtagctgcaagacacctagtcaccttaagtgttcg 1745  
Db 699 ---AAATGGGTGCAAGTTGTATTGGGCTGTTAATGACCTAAGCCATGTTTCGTGGTTTG 643  
Qy 1746 aaacagagctacaatggcttaagggagaaatctgtgaagtctctgtcatctacacgggt 1805  
Db 642 AGGAAGAATTGAATGGCTTCTTTCCAGAGTGTCAAAATCACCATTTTTGTGCACCT--T 586  
Qy 1806 catcagtgaggagatacaaaactcagatgagtcacctaagggtttctcagacaaggagaat 1865  
Db 585 CACCAAAAGAACTCGAAGAGACAGAGGTTCCTCATTTAAACACACACAATAAGCTATTAGAAA 526  
Qy 1866 ctgaaatcacctagatgtagccttaacagagggccagacctcaagagctagtagatcag 1925  
Db 525 TT---CTGGATGTGCGCAGCTTATTGTTGCTAGGCTGATTTGAGAGCTTTTTCAGAGA 469  
Qy 1926 agatcaaatgtcagaactcggagaaacaacatcacttctactcatcgagaccagcga 1985  
Db 468 AGGTCAAAGAACTTCAGATCAATCTAAGATCTAACTTTTCATCAGCTGTGGTCTTCTTA 409  
Qy 1986 ctttcaatgacacttttaggaatcagttgtacaaaggtatcgattcttagtctgaagatag 2045  
Db 408 CATTCAACATAGATTTCAGAGCATCGGTATCGAACAACCTTCGCCAAAAG--CAGAGCATCG 350  
Qy 2046 atgtcgaactagagaggagagatttttacttgtaa 2080  
Db 349 ATGTTGCTCTCAAGAAGAAAGCTTTCGTCTGTTGA 315

[illegible]



BASE COUNT	241 a	263 c	273 g	232 t	6 others
ORIGIN					
Query Match	2.6%; Score 53.4; DB 13; Length 1015;				
Best Local Similarity	46.9%; Pred. No. 0.00091;				
Matches 196; Conservative	2; Mismatches 214; Indels 6; Gaps 1;				
QY	148	gtcaagcttctgactgcaagaacatacaattggctcggaagcgtcaactgcatgcgctta	207		
Db	445	GGCATTACCGTTGTCGGTGGCGCAACGTGAACCTGGCTCGGCACCATCAACCAATTGTATCGC	504		
QY	208	tgagaactcacaatacacaagactctggaactccgcttgatgaactggaagcagcaatg	267		
Db	505	AAACAACACCAAGAGCTGCGGCTCAGGAACACACGGGTTCGGCAGCTTCGCGGAGGTG	564		
QY	268	ctcaagttatcaaggtt-----tacacactggaggacatgaagaacatctacttaatgc	321		
Db	565	CTACGAGAAAGGGGATTTCAACTACACTCTGCGGCACATGACGCGTCTACGCTAACGG	624		
QY	322	aagtaactactctcgctctctgagaataccgataagaagacagttgtttcaaacaggtt	381		
Db	625	CACAGGGTACCTCAGAGACCCACCACAGCGGACCTGTGTAGCCCGGTTTTCACACAGCT	684		
QY	382	gatggcaaatgagcggcctatcactactactatgaggaactatgggatccacttga	441		
Db	685	TCGSTGANACACAGCGGAGTTCGCTGTGCTACTACAAGAGATGAAGACCTCAGCTTCTC	744		
QY	442	tttgatgcatcgaatggtgcgcgtggcgctctctctctctctctctctctctctac	501		
Db	745	TGTGACGCGCAGTCACTGCTGGTGGCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG	804		
QY	502	cgcgcgaactatctgaacatctcaaacgcgtattcggcaagaacattatggcaaat	559		
Db	805	CGTACCACGTTGTTCAATTTCAACGAAAGTTTCTTGGGGCTGAGCTCACCAGCAACT	862		
RESULT	6				
CNS06WLF	958 bp DNA GSS 06-JUL-2001				
LOCUS	T7 end of clone AX0AA027B12 of library AX0AA from strain CBS 7064				
DEFINITION	of Pichia farinosa, genomic survey sequence.				
ACCESSION	AL417865				
VERSION	AL417865.1 GI:12200119				
KEYWORDS	GSS.				
SOURCE	Pichia farinosa.				
ORGANISM	Pichia farinosa				
REFERENCE	Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;				
AUTHORS	Pichia farinosa				
TITLE	Saccharomycetales; Saccharomycetaceae; Pichia.				
JOURNAL	1 (bases 1 to 958)				
MEDLINE	de Montigny, J., Spohner, C., Souciet, J., Tekala, F., Dujon, B.,				
REFERENCE	Wincker, P., Artiguenave, F. and Potier, S.				
AUTHORS	Genomic exploration of the hemiascomycetous yeasts: 15. Pichia				
TITLE	sorbitophila				
JOURNAL	FEMS Lett. 487 (1), 87-90 (2000)				
MEDLINE	20584725				
REFERENCE	2 (bases 1 to 958)				
AUTHORS	Souciet, J.L., Aigle, M., Artiguenave, F., Blandin, G.,				
TITLE	Bolotin-Fukuhara, M., Bon, E., Brothier, P., Casaregola, S.,				
JOURNAL	de-Montigny, J., Dujon, B., Durrens, P., Lepingle, A., Llorente, B.,				
MEDLINE	Malpertuy, A., Neuveglise, C., Ozier-Kalogeropoulos, O., Potier, S.,				
REFERENCE	Saurin, W., Tekala, F., Toffano-Nioche, C., Wesolowski-Louvel, M.,				
AUTHORS	Wincker, P. and Weissenbach, J.				
TITLE	Genomic exploration of the hemiascomycetous yeasts: 1. A set of				
JOURNAL	yeast species for molecular evolution studies				
MEDLINE	FEMS Lett. 487 (1), 3-12 (2000)				
REFERENCE	20584711				
AUTHORS	3 (bases 1 to 958)				
TITLE	Genoscope.				
JOURNAL	Direct Submission				
MEDLINE	Submitted (08-SEP-2000) Genoscope - Centre National de Sequencage,				
REFERENCE	2 rue Gaston Cremieux, CP 5706, 91057 EVRY cedex, FRANCE. (E-mail :				
AUTHORS	seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)				
TITLE	JOURNAL				
MEDLINE	JOURNAL				
REFERENCE	JOURNAL				
AUTHORS	JOURNAL				
TITLE	JOURNAL				
MEDLINE	JOURNAL				
REFERENCE	JOURNAL				
AUTHORS	JOURNAL				
TITLE	JOURNAL				
MEDLINE	JOURNAL				
REFERENCE	JOURNAL				
AUTHORS	JOURNAL				
TITLE	JOURNAL				
MEDLINE	JOURNAL				
REFERENCE	JOURNAL				
AUTHORS	JOURNAL				
TITLE	JOURNAL				
MEDLINE	JOURNAL				
REFERENCE	JOURNAL				
AUTHORS	JOURNAL				
TITLE	JOURNAL				
MEDLINE	JOURNAL				
REFERENCE	JOURNAL				
AUTHORS	JOURNAL				
TITLE	JOURNAL				
MEDLINE	JOURNAL				
REFERENCE	JOURNAL				
AUTHORS	JOURNAL				
TITLE	JOURNAL				
MEDLINE	JOURNAL				
REFERENCE	JOURNAL				
AUTHORS	JOURNAL				
TITLE	JOURNAL				
MEDLINE	JOURNAL				
REFERENCE	JOURNAL				
AUTHORS	JOURNAL				
TITLE	JOURNAL				
MEDLINE	JOURNAL				
REFERENCE	JOURNAL				
AUTHORS	JOURNAL				
TITLE	JOURNAL				
MEDLINE	JOURNAL				
REFERENCE	JOURNAL				
AUTHORS	JOURNAL				
TITLE	JOURNAL				
MEDLINE	JOURNAL				
REFERENCE	JOURNAL				
AUTHORS	JOURNAL				
TITLE	JOURNAL				
MEDLINE	JOURNAL				
REFERENCE	JOURNAL				
AUTHORS	JOURNAL				
TITLE	JOURNAL				
MEDLINE	JOURNAL				
REFERENCE	JOURNAL				
AUTHORS	JOURNAL				
TITLE	JOURNAL				
MEDLINE	JOURNAL				
REFERENCE	JOURNAL				
AUTHORS	JOURNAL				
TITLE	JOURNAL				
MEDLINE	JOURNAL				
REFERENCE	JOURNAL				
AUTHORS	JOURNAL				
TITLE	JOURNAL				
MEDLINE	JOURNAL				
REFERENCE	JOURNAL				
AUTHORS	JOURNAL				
TITLE	JOURNAL				
MEDLINE	JOURNAL				
REFERENCE	JOURNAL				
AUTHORS	JOURNAL				
TITLE	JOURNAL				
MEDLINE	JOURNAL				
REFERENCE	JOURNAL				
AUTHORS	JOURNAL				
TITLE	JOURNAL				
MEDLINE	JOURNAL				
REFERENCE	JOURNAL				
AUTHORS	JOURNAL				
TITLE	JOURNAL				
MEDLINE	JOURNAL				
REFERENCE	JOURNAL				
AUTHORS	JOURNAL				
TITLE	JOURNAL				
MEDLINE	JOURNAL				
REFERENCE	JOURNAL				
AUTHORS	JOURNAL				
TITLE	JOURNAL				
MEDLINE	JOURNAL				
REFERENCE	JOURNAL				
AUTHORS	JOURNAL				
TITLE	JOURNAL				
MEDLINE	JOURNAL				
REFERENCE	JOURNAL				
AUTHORS	JOURNAL				
TITLE	JOURNAL				
MEDLINE	JOURNAL				
REFERENCE	JOURNAL				
AUTHORS	JOURNAL				
TITLE	JOURNAL				
MEDLINE	JOURNAL				
REFERENCE	JOURNAL				
AUTHORS	JOURNAL				
TITLE	JOURNAL				
MEDLINE	JOURNAL				
REFERENCE	JOURNAL				
AUTHORS	JOURNAL				
TITLE	JOURNAL				
MEDLINE	JOURNAL				
REFERENCE	JOURNAL				
AUTHORS	JOURNAL				
TITLE	JOURNAL				
MEDLINE	JOURNAL				
REFERENCE	JOURNAL				
AUTHORS	JOURNAL				
TITLE	JOURNAL				
MEDLINE	JOURNAL				
REFERENCE	JOURNAL				
AUTHORS	JOURNAL				
TITLE	JOURNAL				
MEDLINE	JOURNAL				
REFERENCE	JOURNAL				
AUTHORS	JOURNAL				
TITLE	JOURNAL				
MEDLINE	JOURNAL				
REFERENCE	JOURNAL				
AUTHORS	JOURNAL				
TITLE	JOURNAL				
MEDLINE	JOURNAL				
REFERENCE	JOURNAL				
AUTHORS	JOURNAL				
TITLE	JOURNAL				
MEDLINE	JOURNAL				
REFERENCE	JOURNAL				
AUTHORS	JOURNAL				
TITLE	JOURNAL				
MEDLINE	JOURNAL				
REFERENCE	JOURNAL				
AUTHORS	JOURNAL				
TITLE	JOURNAL				
MEDLINE	JOURNAL				
REFERENCE	JOURNAL				
AUTHORS	JOURNAL				
TITLE	JOURNAL				
MEDLINE	JOURNAL				
REFERENCE	JOURNAL				
AUTHORS	JOURNAL				
TITLE	JOURNAL				
MEDLINE	JOURNAL				
REFERENCE	JOURNAL				
AUTHORS	JOURNAL				
TITLE	JOURNAL				
MEDLINE	JOURNAL				
REFERENCE	JOURNAL				
AUTHORS	JOURNAL				
TITLE	JOURNAL				
MEDLINE	JOURNAL				
REFERENCE	JOURNAL				
AUTHORS	JOURNAL				
TITLE	JOURNAL				
MEDLINE	JOURNAL				
REFERENCE	JOURNAL				
AUTHORS	JOURNAL				
TITLE	JOURNAL				
MEDLINE	JOURNAL				
REFERENCE	JOURNAL				
AUTHORS	JOURNAL				
TITLE	JOURNAL				
MEDLINE	JOURNAL				
REFERENCE	JOURNAL				
AUTHORS	JOURNAL				
TITLE	JOURNAL				
MEDLINE	JOURNAL				
REFERENCE	JOURNAL				
AUTHORS	JOURNAL				
TITLE	JOURNAL				
MEDLINE	JOURNAL				
REFERENCE	JOURNAL				
AUTHORS	JOURNAL				
TITLE	JOURNAL				
MEDLINE	JOURNAL				
REFERENCE	JOURNAL				
AUTHORS	JOURNAL				
TITLE	JOURNAL				
MEDLINE	JOURNAL				
REFERENCE	JOURNAL				
AUTHORS	JOURNAL				
TITLE	JOURNAL				
MEDLINE	JOURNAL				
REFERENCE	JOURNAL				
AUTHORS	JOURNAL				
TITLE	JOURNAL				
MEDLINE	JOURNAL				
REFERENCE	JOURNAL				
AUTHORS	JOURNAL				
TITLE	JOURNAL				
MEDLINE	JOURNAL				
REFERENCE	JOURNAL				
AUTHORS	JOURNAL				
TITLE	JOURNAL				
MEDLINE	JOURNAL				
REFERENCE	JOURNAL				
AUTHORS	JOURNAL				
TITLE	JOURNAL				
MEDLINE	JOURNAL				
REFERENCE	JOURNAL				
AUTHORS	JOURNAL				
TITLE	JOURNAL				
MEDLINE	JOURNAL				
REFERENCE	JOURNAL				
AUTHORS	JOURNAL				
TITLE	JOURNAL				
MEDLINE	JOURNAL				
REFERENCE	JOURNAL				
AUTHORS	JOURNAL				
TITLE	JOURNAL				
MEDLINE	JOURNAL				
REFERENCE	JOURNAL				
AUTHORS	JOURNAL				
TITLE	JOURNAL				
MEDLINE	JOURNAL				
REFERENCE	JOURNAL				
AUTHORS	JOURNAL				
TITLE	JOURNAL				
MEDLINE	JOURNAL				
REFERENCE	JOURNAL				
AUTHORS	JOURNAL				
TITLE	JOURNAL				
MEDLINE	JOURNAL				
REFERENCE	JOURNAL				
AUTHORS	JOURNAL				
TITLE					









```
BASE COUNT      201 a      64 c      131 g      202 t      503 others
ORIGIN

Query Match      2.0%; Score 42.8; DB 13; Length 1101;
Best Local Similarity 17.3%; Pred. No. 0.82;
Matches 109; Conservative 252; Mismatches 264; Indels 5; Gaps 2;

QY 545 acattatggaacattctgttaagaagctcttattaccacaaagcattacaaagactac 604
Db 416 AWAATAAATAATTTWAAWAAWAAATAATTTWAAWAAWAAWAAWAAWAAWAAWAAWAA 475
QY 605 aacgagagaaacttctatctcttggaacgtttgccattcaactttcaaacctcaggcaaa 664
Db 476 AAAAAAATTTTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWT 535
QY 665 ggaactgtagtcttattcttctgactgactattctctcactctctcttcggagacataa 724
Db 536 AWATAAATTTTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWT 595
QY 725 atcaagtgcacacattctacgatacctagatgagagaagatacaatggcattcgtctca 784
Db 596 TMTTHAWAHTTWTHTHTYAWWHTWTHTWTWAAWHTTWTAAAYYYTCMYHHYHHWH 655
QY 785 cgccgtgactgactggaacgcctcttctcccggtggtgtacatttctcggtaccgg 844
Db 656 HAHAAWAAWTTTWTWTHAYHWATYHYYYMYCAMCMCTTCHHCYYYHHVTAHTHT 715
QY 845 acaaccccttaccacaaacacaggaattgagcttta--gtaacttcaacttttaccaca 902
Db 716 HHNYAHYMYMYWYMYWYMYWYMYWYMYWYMYWYMYWYMYWYMYWYMYWYMYWYMYW 775
QY 903 aatggcagcatcgctctctctctgtagcgtcgctccattcaactgttatgaccgctt 962
Db 776 AAWAAWATHTHTHTHTHTHTHTHTHTHTHTHTHTHTHTHTHTHTHTHTHTHTHT 835
QY 963 caggagtgaaac---gaggagattccagctctcttgaagaattcttaactcagatggg 1019
Db 836 WTHWYHTHTHTHTHTHTHTHTHTHTHTHTHTHTHTHTHTHTHTHTHTHTHTHT 895
QY 1020 gaatagtagcacaactctctatgtccatcatcattctccagtcgagagaaggtctcagga 1079
Db 896 ACHAHHTHTCMCHHHMHCTCHHHHTHTWHTCTHWHWHHHWHHHWHWATWHTTWTMM 955
QY 1080 accgaggttatgaaattcttcttattctacaaagcattgaaacatcatgtttatcatag 1139
Db 956 CCMHHCHYHHYHHYHHYHHYHHYHHYHHYHHYHHYHHYHHYHHYHHYHHYHHYHH 1015
1140 ctatgtattaccattgccacacacataggat 1169
Db 1016 HAMTTATWMMHWHWAAHWATWMMWMMWAT 1045

RESULT 14
CNS07DHO      1043 bp      DNA      GSS      08-JUL-2001
LOCUS
DEFINITION
T3 end of clone BD0AA012G01 of library BD0AA from strain CBS 94 of
Candida tropicalis, genomic survey sequence.
ACCESSION
AL440482
VERSION
AL440482.1 GI:12223893
KEYWORDS
GSS.
SOURCE
Candida tropicalis.
ORGANISM
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
Saccharomycetales; microsporid Saccharomycetales; Candida.
REFERENCE
1. (bases 1 to 1043)
AUTHORS
Blandin,G., Ozler-Kalogeropoulos,O., Wincker,P., Artiguenave,F. and
Dujon,B.
TITLE
Genomic exploration of the hemiascomycetous yeasts: 16. Candida
tropicalis
JOURNAL
FEBS Lett. 487 (1), 91-94 (2000)
MEDLINE
20584726
REFERENCE
2 (bases 1 to 1043)

AUTHORS
Souciet,J.L., Aigle,M., Artiguenave,F., Blandin,G.,
Bolotin-Fukuhara,M., Bon,E., Brottier,P., Casaregola,S.,
de-Montigny,J., Dujon,B., Durrens,P., Lepingle,A., Llorente,B.,
Malpertuy,A., Neuveglise,C., Ozler-Kalogeropoulos,O., Potier,S.,
Saurin,W., Tekia,F., Toffano-Nioche,C., Wesolowski-Louvel,M.,
Wincker,P. and Weissenbach,J.
TITLE
Genomic exploration of the hemiascomycetous yeasts: 1. A set of
yeast species for molecular evolution studies
JOURNAL
FEBS Lett. 487 (1), 3-12 (2000)
MEDLINE
20584711
REFERENCE
3 (bases 1 to 1043)
AUTHORS
Genoscope.
DIRECT SUBMISSION
Submitted (08-SEP-2000) Genoscope - Centre National de Sequencage,
2 rue Gaston Crémieux, CP 5706, 91057 EVRY cedex, FRANCE. (E-mail :
seqref@genoscope.cns.fr - Web :
http://genoscope.cns.fr)
COMMENT
This GSS is part of a random genomic sequencing program of thirteen
yeast species: Saccharomyces bayanus var. uvarum, Saccharomyces
exiguus, Saccharomyces servazzii, Zygosaccharomyces rouxii,
Saccharomyces kluyveri, Kluyveromyces thermotolerans, Kluyveromyces
lactis var. lactis, Kluyveromyces marxianus var. marxianus, Pichia
angusta, Debaryomyces hansenii var. hansenii, Pichia sorbitophila,
Candida tropicalis and varrowia lipolytica. Genomic inserts of 3 to
5 kb were prepared and both extremities were sequenced. See
keywords for description of this sequence and for the sequence of
the other extremity of this insert.
FEATURES
Location/Qualifiers
1..1043
/organism="Candida tropicalis"
/strain="CBS 94"
/db_xref="taxon:5482"
/clone="BD0AA012G01"
/clone_lib="BD0AA"
/notes="end : T3"
misc_feature
<2..>985
/notes="similar to Saccharomyces cerevisiae ORF YKL220c [
FRE2; ferric (and cupric) reductase ]
similar to Saccharomyces cerevisiae ORF YNR060w [ FRE4 ;
strong similarity to Fre2p, and similarity to Fre1p ]
similar to Saccharomyces cerevisiae ORF YOR381w [ FRE3 ;
strong similarity to ferric reductase Fre2p ]
similar to Saccharomyces cerevisiae ORF YOR384w [ FRE5 ;
strong similarity to ferric reductase Fre2p ]"
/evidence="not_experimental"
BASE COUNT      287 a      173 c      235 g      345 t
ORIGIN

Query Match      2.0%; Score 41; DB 13; Length 1043;
Best Local Similarity 52.0%; Pred. No. 2.5;
Matches 92; Conservative 0; Mismatches 85; Indels 0; Gaps 0;

QY 1072 cttcagaacccaggttatgaaattcttcttattcaccacaaagcattgacatgatt 1131
Db 1 CTTCAGAAGACGATGGTATGAAATGTTTGTTCATACACATCGTTATGGTCGCATTTA 60
QY 1132 tatcatgactatgtattaccattgccacacacattgagtgatggctggatctggtccat 1191
Db 61 TATTGTGGTACGTGGATCCATGTTGATGATCTTGGATATCTTTGGTTCTCTAGCCAG 120
QY 1192 ggctggcatctctgcttcgacaggttctgcggaattgtacgtatcatcagaacgg 1248
Db 121 TTTTGCACCTTGGTGTGTTTGTATGATTGATTGAGGTTGTTCTGTTGG 177

RESULT 15
BG714799
LOCUS
DEFINITION
602677117F1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:4799501 5',
mRNA sequence.
ACCESSION
BG714799
VERSION
BG714799.1 GI:13993730
KEYWORDS
EST.
```

SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 2149)  
REFERENCE NIH-MGC <http://mgs.nci.nih.gov/>.  
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)  
TITLE Unpublished (1999)  
JOURNAL  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgaphs-remail.nih.gov  
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.  
cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki  
Toshiyuki and Piero Carninci (RIKEN)  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: L1AM10689 row: a column: 06  
High quality sequence start: 10  
High quality sequence stop: 22.  
Location/Qualifiers  
1..2149  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4799501"  
/clone\_lib="NIH\_MGC\_96"  
/tissue\_type="hypothalamus"  
/lab\_host="DH10B"  
/notes="Organ: brain; Vector: pBluescriptR (modified  
pBluescript KS+); Site\_1: BamHI; Site\_2: SalI-XhoI (gtcggag  
); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTVN-3',  
size-selected for average insert size 2.3 kb and  
normalized to ROT 5. This is a primary library enriched  
for full-length clones and constructed using the  
Cap-trapper method (Carninci, in preparation). Library  
constructed by M. Brownstein (NIH/NHGRI, National  
Institutes of Health). Note: this is a NIH\_MGC Library."  
BASE-COUNT 616 a 871 c 192 g 470 t  
ORIGIN

Query Match 2.0%; Score 41; DB 11; Length 2149;  
Best Local Similarity 51.4%; Pred. No. 3.3;  
Matches 95; Conservative 0; Mismatches 90; Indels 0; Gaps 0;  
Qy 1295 atcaagatctctgtcaagaagcctaagtcttctcaagtatcaagtggagcatttgcctat 1354  
1281 ATCCCATATCCAGCCCATACACTAATTCATCATCAATACACTCGAATCTTACCTAT 1340  
Qy 1355 atgtacttcttccacaaatcagcctgggtctacagtttctcaatctcatccctcaca 1414  
Db 1341 ACACATTACCTCAATACATACATCCCACTACCACTACCACTTCAACTCAATCTTACCTCT 1400  
Qy 1415 gtctctcagaagagcagagatcctcaaacaccagatcaactaactatgtacgtcaaa 1474  
Db 1401 GATCAATAAACATACATACATAGCCCACTCAATATCAACTAACCTTCTTACCACTAA 1460  
Qy 1475 gctaa 1479  
Db 1461 AATTA 1465

Search completed: May 14, 2002, 06:33:23  
Job time: 8015 sec

